

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:55:26 ; Search time 21 seconds
(without alignments)
854.276 Million cell updates/sec

Title: US-10-016-985-2
Perfect score: 2239
Sequence: 1 MSVGCPEPPRSLTCCGPG.....GRADSKGQVVLATAICV 424

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pdp.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pdp.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pdp.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pdp.*
- 5: /cgn2_6/ptodata/1/iaa/ECTUS.COMB.pdp.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	100.0	424	4	US-09-816-094-2
2	2084.5	93.1	417	4	US-09-816-094-4
3	341.5	15.3	729	2	US-08-677-298-2
4	341.5	15.3	729	4	US-09-523-849-33
5	340	15.2	513	3	US-08-924-183-8
6	340	15.2	513	4	US-09-488-364-8
7	336	15.0	765	4	US-09-975-326-4
8	335.5	15.0	766	4	US-09-975-326-2
9	334	14.9	793	4	US-09-523-849-32
10	333.5	14.9	779	4	US-08-817-832B-31
11	333	14.9	476	3	US-08-924-183-1
12	333	14.9	476	4	US-09-488-364-2
13	328.5	14.7	448	2	US-09-159-385-2
14	328.5	14.7	448	3	US-09-186-277-2
15	328	14.6	476	3	US-09-522-800-16
16	328	14.6	476	3	US-08-924-183-1
17	328	14.6	476	4	US-09-488-364-1
18	328	14.6	476	4	US-09-113-785-1
19	323.5	14.4	722	4	US-09-984-890-4
20	323.5	14.4	722	4	US-08-817-832B-32
21	322.5	14.4	483	3	US-08-924-183-9
22	322.5	14.4	483	4	US-09-488-364-9
23	322.5	14.4	745	4	US-09-523-849-36
24	321.5	14.4	724	4	US-09-984-890-2
25	318	14.2	454	2	US-09-159-385-1
26	318	14.2	454	3	US-09-186-277-1
27	315	14.1	552	3	US-08-557-006C-40

28	311.5	13.9	582	4	US-09-428-711A-2	Sequence 2, Appli
29	311.5	13.9	1051	4	US-09-428-711A-14	Sequence 14, Appli
30	310	13.8	668	4	US-09-930-181-2	Sequence 2, Appli
31	309	13.8	633	3	US-08-557-006C-43	Sequence 43, Appli
32	308.5	13.8	257	2	US-07-857-224B-25	Sequence 25, Appli
33	307	13.7	543	4	US-09-529-093A-2	Sequence 2, Appli
34	307	13.7	543	4	US-09-529-154-2	Sequence 2, Appli
35	305.5	13.6	476	3	US-08-924-183-7	Sequence 7, Appli
36	305.5	13.6	476	3	US-09-488-364-7	Sequence 7, Appli
37	302	13.5	631	4	US-09-579-664B-11	Sequence 11, Appli
38	302	13.5	776	4	US-09-523-849-34	Sequence 34, Appli
39	298.5	13.3	345	3	US-09-101-146-1	Sequence 1, Appli
40	298	13.3	604	4	US-09-523-849-35	Sequence 35, Appli
41	295.5	13.2	257	3	US-09-101-146-6	Sequence 6, Appli
42	294.5	13.2	396	4	US-09-801-876B-2	Sequence 2, Appli
43	294.5	13.2	1050	4	US-09-428-711A-16	Sequence 16, Appli
44	294	13.1	436	4	US-09-734-673-2	Sequence 2, Appli
45	294	13.1	436	4	US-09-523-849-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-816-094-2

; Sequence 2, Application US/09816094

; Patent No. 6534299

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: THEREOF

; FILE REFERENCE: CL000536-CIP

; CURRENT APPLICATION NUMBER: US/09/816,094

; CURRENT FILING DATE: 2001-03-26

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 424

; TYPE: PRT

; ORGANISM: Human

US-09-816-094-2

Query Match

Best Local Similarity 100.0%; Score 2239; DB 4; Length 424;

Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSVGCPEPPRSLTCCGPGTAPGAGVPLLTEDMOALTTLTAAADVTVKHYELVRELG	60
Db	1	MSVGCPEPPRSLTCCGPGTAPGAGVPLLTEDMOALTTLTAAADVTVKHYELVRELG	60
Qy	61	KGTGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSTNSLSSPFFIKKVDVVFET	120
Db	61	KGTGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSTNSLSSPFFIKKVDVVFET	120
Qy	121	EDCVFAQYAPAGDLFDIIPPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKENV	180
Db	121	EDCVFAQYAPAGDLFDIIPPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKENV	180
Qy	181	LLFDRECRVKLADFGMTTRVGRVKRVSGTIPTTAPVCOAGRADGLAVDTGVDVWAFG	240
Db	181	LLFDRECRVKLADFGMTTRVGRVKRVSGTIPTTAPVCOAGRADGLAVDTGVDVWAFG	240
Qy	241	VLFVLTGTFNPPNAAAGADAFEEFVVRWQGRGLPGLPSOWRRFTEPALRMFORLLALEP	300
Db	241	VLFVLTGTFNPPNAAAGADAFEEFVVRWQGRGLPGLPSOWRRFTEPALRMFORLLALEP	300
Qy	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPCDRPPAAGPLRLEAPGLKRTVLTESG	360
Db	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPCDRPPAAGPLRLEAPGLKRTVLTESG	360
Qy	361	SGSRPAPPVGVSLPVPVVPVVPVPEPGLAPGPPGRTDGRADSKGQVVLATAI	420

5/17/02
60205228

RESULT 4
US-09-523-849-33
; Sequence 33, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:55:26 ; Search time 21 Seconds
(without alignments)
854.276 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239
Sequence: 1 MSVGCPEPEPRSLTCCGPG.....GRADKSKGVVLATAICV 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	2239	100.0	424	4	US-09-816-094-2	Sequence 2, Appli
2	2084.5	93.1	417	4	US-09-816-094-4	Sequence 4, Appli
3	341.5	15.3	729	2	US-08-677-298-2	Sequence 2, Appli
4	341.5	15.3	729	4	US-09-523-849-33	Sequence 33, Appli
5	340	15.2	513	3	US-08-924-183-8	Sequence 8, Appli
6	340	15.2	513	4	US-09-488-364-8	Sequence 8, Appli
7	336	15.0	765	4	US-09-975-326-4	Sequence 4, Appli
8	335.5	15.0	765	4	US-09-975-326-2	Sequence 2, Appli
9	334	14.9	793	4	US-09-523-849-32	Sequence 32, Appli
10	333.5	14.9	779	4	US-08-817-832B-31	Sequence 31, Appli
11	333	14.9	476	3	US-08-924-183-2	Sequence 2, Appli
12	333	14.9	476	4	US-09-488-364-2	Sequence 2, Appli
13	328.5	14.7	448	2	US-09-159-385-2	Sequence 2, Appli
14	328.5	14.7	448	3	US-09-186-277-2	Sequence 2, Appli
15	328	14.6	476	3	US-08-522-800-16	Sequence 16, Appli
16	328	14.6	476	3	US-08-924-183-1	Sequence 1, Appli
17	328	14.6	476	4	US-09-488-364-1	Sequence 1, Appli
18	328	14.6	476	4	US-09-113-785-1	Sequence 1, Appli
19	323.5	14.4	722	4	US-09-984-890-4	Sequence 4, Appli
20	323.5	14.4	722	4	US-08-817-832B-32	Sequence 32, Appli
21	322.5	14.4	483	3	US-08-924-183-9	Sequence 9, Appli
22	322.5	14.4	483	4	US-09-488-364-9	Sequence 9, Appli
23	322.5	14.4	745	4	US-09-523-849-36	Sequence 36, Appli
24	321.5	14.4	724	4	US-09-984-890-2	Sequence 2, Appli
25	318	14.2	454	2	US-09-159-385-1	Sequence 1, Appli
26	318	14.2	454	3	US-09-186-277-1	Sequence 1, Appli
27	315	14.1	552	3	US-08-557-006C-40	Sequence 40, Appli

RESULT 1

US-09-816-094-2

; Sequence 2, Application US/09816094

; Patent No. 6534299

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: THERBOF

; CURRENT FILING DATE: 2001-03-26

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 424

; TYPE: PRT

; ORGANISM: Human

US-09-816-094-2

Query Match	Best Local Similarity	Score	2239; DB 4;	Length 424;
Matches	424;	Conservative	0;	Mismatches
Qy	1	MSVGCPEPEPRSLTCCGPGTAPGPGAGVPLLTEDMOALTIRTLAASDVTKHVELREL	60	
Db	1	MSVGCPEPEPRSLTCCGPGTAPGPGAGVPLLTEDMOALTIRTLAASDVTKHVELREL	60	
Qy	61	KGTGKVDLVVYKGTGTMKALKFNKSKTKLKNFLREVSTNSLSSPFIKVFDDVVFET	120	
Db	61	KGTGKVDLVVYKGTGTMKALKFNKSKTKLKNFLREVSTNSLSSPFIKVFDDVVFET	120	
Qy	121	EDCVFAQAEYAPAGDLFDIIPQVGLPEDTKRCVQQLGLALDFMHGRQLVHRDIKPEV	180	
Db	121	EDCVFAQAEYAPAGDLFDIIPQVGLPEDTKRCVQQLGLALDFMHGRQLVHRDIKPEV	180	
Qy	181	LLFDRECRVKLADFGMTREVCGRVKVSCTIPTYTAPEVCOAGRADGLAVDTGVDVWAFG	240	
Db	181	LLFDRECRVKLADFGMTREVCGRVKVSCTIPTYTAPEVCOAGRADGLAVDTGVDVWAFG	240	
Qy	241	VLIFCVLTGNFPPWEAASGADAFEEFVRWQGRGLPGLPSQWRFFTEPALMFQRLALEP	300	
Db	241	VLIFCVLTGNFPPWEAASGADAFEEFVRWQGRGLPGLPSQWRFFTEPALMFQRLALEP	300	
Qy	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG	360	
Db	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG	360	
Qy	361	SGSRPAPPAVGSVPLVPVFPVVPVPEPGLACQGPRTDGRADKSKGVVLATAI	420	

5/17/09
60205228

Db 60 IVKLYD-VFEDDDKLYVMEYCYGSGDLFDLKKRGLSEDEARFYARQILSALEYLHSQG 118
QY 170 LVHRDKPENVLFFDRECRVVKLADFGMTTRV--GCRVKRVSGTTPYTAPEVCQAGRAD 256
Db 119 IHRDLKPENILL-DSD-GHVKLADGLAKQLDSGGTLLTTFVTGYVAPEVL-LGKG 175
QY 227 GLAVDTGVDVWAFGVILFCVLTGNFPWEAASGADAFPEEFVRWQGRGLPGLPSQWRRT 285
Db 176 GKX---VDIWSGLVILYELLTGKPPF---PGDDQLLALFKKIGKPPPPPPPEW-KISP 226
QY 287 PALRMFQRLALALEPERRGPAKE 308
Db 228 EAKDLIKKLLVXDPKRLTAEE 249

RESULT 15
US-09-976-782-30
; Sequence 30, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,669
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/252,455
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/240,648
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Consensus
; OTHER INFORMATION: sequence
US-09-976-782-30

Query Match 15.6%; Score 350; DB 12; Length 254;
Best Local Similarity 35.9%; Pred. No. 7.4e-18;
Matches 97; Conservative 44; Mismatches 105; Indels 24; Gaps 10;
QY 53 YELVRELGKGYGVKVDLVVYKGTGKALKXFNKS--KTKLKNPLREVSITNSLSSPFI 110
Db 1 YELGKLGSGAGFYGVKGHKDTGEIVAKITKKRSJSEKKRFLREIQLRL-SHPNI 59
QY 111 IKVFDVVPETEDCVFAQEVAPAGLDIIPPQ-VGLPDTVVKRCVQQLGLALDFMEGRQ 169
Db 60 VRLLG-VFEEDHLYVMEYMEGGDLFDYLRNGLLSEKAKKIALQILRGLEYLHSG 118
QY 170 LVHRDKPENVLFFDRECRVVKLADFGMTTRVGC-----RVKRVSGTTPYTAPEVCQAGRA 225
Db 119 IVHRDLKPENILL--DENGTVKIADFGIARKLESSSYEKLATTFVGTPEYMAPEV-----L 171

QY 226 DGLAVDTGVDVWAFGVILFCVLTGNFPWEAASGADAFPEEFVRWQGRGLPGLPSQWRRT 285
Db 172 EGRGYSSKVDVMSGLVILYELLTGKLPFPIDPLEBLPRIKER-PRRLPLLPN-----CS 226
QY 286 EPALRMFQRLALALEPERRGPAKEVREFLKH 315
Db 227 BELKDLIKKLLXNDPEKPTAKEI---LNH 253

Search completed: December 5, 2003, 10:05:22
Job time : 39 secs


```
RESULT 11
US-10-353-690-50
; Sequence 50, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodini, Jacqueline
; APPLICANT: Rodrigue-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69232, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MPI02-018P1NOMXIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-353-690-50

Query Match 24.4%; Score 546.5; DB 12; Length 348;
Best Local Similarity 42.9%; Pred. No. 5e-32;
Matches 117; Conservative 42; Mismatches 111; Indels 3; Gaps 3;

QY 34 EDMQALTRTLAASDVTKHYELVRELKGTGKVDLVVYKGTGKWKALKFVNKSKTKLN 93
DB 43 EDMWTLAQTLVRAEVDLEYEVRLGQGRYGRVLLVTHRQKGTPLALKQPKPRTSRG 102
QY 94 FLREVSITNLSSSPFIKFDVWFETEDCVYFAQYAPAGDLFDIIPQVGLPDPDTVKR 153
DB 103 FLYEFCVGLSLGAHSAIVTAYGIGIESAHSYFLTEPVLHGLDMAFIQPKVGLPQPAVHR 162
QY 154 CVQQLGLALDFMHGQLVHRDKIPENVLLFDECRVRKLAQFGMTVRVRCVRKRVSGTIP 213
DB 163 CAAQLASALEYTHARGLVRLDKPENLVLCVPACRRFKLTDFGHTRPRGTLRLLAGPPIP 222
QY 214 YTAPEVC-QAGRADGLAVDTGVDVWAFGLVLCVLTGNFPNE-AAGSADAFEEFVRVQR 271

RESULT 12
US-09-863-776-40
; Sequence 40, Application US/09863776
; Publication No. US20030198953A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spaderna, Steven K
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Li, Li
; APPLICANT: Taupier, Raymond J
; APPLICANT: Gangolli, Esha
; TITLE OF INVENTION: No. US20030198953A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-020
; CURRENT APPLICATION NUMBER: US/09/863,776
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/206,679
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/206,688
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/206,829
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 60/207,748
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/207,798
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/208,263
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/208,831
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/209,451
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: 60/210,060
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: 60/219,507
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/221,337
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/221,927
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 60/263,135
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,688
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/263,694
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Serine/threonine protein kinase, catalytic domain
; OTHER INFORMATION: Consensus Sequence
US-09-863-776-40

Query Match 17.6%; Score 394; DB 12; Length 256;
```

Db 181 LLEDECRVVKLADFGMTVRVCRVRSVGTIPYTAPEVCQAGRADGFAVDTGVVDVWAFG 240
Qy 241 VLIFCVLTGNFPWEAASGADAFFEFVVRWQGRPLGFLPSQWRRFTEPALRMFORLLALEP 300
Db 241 VLIFCVLTGNFPWEAASGADAFFEFVVRWQGRPLGFLPSQWRRFTEPALRMFORLLALEP 300
Qy 301 ERGGPAKVVFLFKHLSLSELRSSHRARPPGDRPPAAGPLRLAEAGPLKRTVLTESG 360
Db 301 ERGGPAKVVFLFKHLSLSELRSSHRARPPGDRPPAAGPLRLAEAGPLKRTVLTESG 358
Qy 361 SGSRPAPVAGSVPLFVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 420
Db 359 SGSRPAPVAGSVPLFVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 413
Qy 421 EICV 424
Db 414 EICV 417

RESULT 9
US-09-916-790-2
; Sequence 2, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Sifos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916, 790
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-916-790-2

Query Match 81.4%; Score 1822; DB 9; Length 346;
Best Local Similarity 100.0%; Pred. No. 4.7e-125;
Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 MALFVNKSKTKLNFLEVSITNSLSPPFIKVFVDFVVFTEDCYVFAQYAPAGLFD 138
Db 1 MALFVNKSKTKLNFLEVSITNSLSPPFIKVFVDFVVFTEDCYVFAQYAPAGLFD 60
Qy 139 IIPQVGLPDTVKRCVQOGLALDFMHGROLVHRDIPENLVLFDECRVVKLADFGMT 198
Db 61 IIPQVGLPDTVKRCVQOGLALDFMHGROLVHRDIPENLVLFDECRVVKLADFGMT 120
Qy 199 RVGCRVRSVGTIPYTAPEVCQAGRADGLAVDTGVVDVWAFGLIFCVLTGNFPWEAASG 258
Db 121 RVGCRVRSVGTIPYTAPEVCQAGRADGLAVDTGVVDVWAFGLIFCVLTGNFPWEAASG 180
Qy 259 ADAFFEFVRWQGRPLGFLPSQWRRFTEPALRMFORLLALEPERRGPAKEVFRFLKHILT 318
Db 181 ADAFFEFVRWQGRPLGFLPSQWRRFTEPALRMFORLLALEPERRGPAKEVFRFLKHILT 240
Qy 319 SELRRSSHRARPPGDRPPAAGPLRLAEAGPLKRTVLTESGSRPAPVAGSVPLFVP 378
Db 241 SELRRSSHRARPPGDRPPAAGPLRLAEAGPLKRTVLTESGSRPAPVAGSVPLFVP 300
Qy 379 VPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 424
Db 301 VPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 346

RESULT 10
US-10-288-798-16

; Sequence 16, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Dannel B;
; APPLICANT: WALIA, Navinder K.; HAPALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARZANI, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dyoung Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUNAR, Javalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil

; TITLE OF INVENTION: HUMAN KINASES

; FILE REFERENCE: PI-0209 USA

; CURRENT APPLICATION NUMBER: US/10/288,798

; CURRENT FILING DATE: 2002-11-01

; PRIOR APPLICATION NUMBER: PCT/US01/27219

; PRIOR FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/240,542

; PRIOR FILING DATE: 2000-10-13

; PRIOR APPLICATION NUMBER: US 60/238,389

; PRIOR FILING DATE: 2000-10-06

; PRIOR APPLICATION NUMBER: US 60/236,499

; PRIOR FILING DATE: 2000-09-29

; PRIOR APPLICATION NUMBER: US 60/234,902

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: US 60/232,654

; PRIOR FILING DATE: 2000-09-14

; PRIOR APPLICATION NUMBER: US 60/231,357

; PRIOR FILING DATE: 2000-09-08

; PRIOR APPLICATION NUMBER: US 60/229,873

; PRIOR FILING DATE: 2000-08-31

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PERL Program

; SEQ ID NO 16

; LENGTH: 348

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030207299A1 7477062CD1

US-10-288-798-16

Query Match 24.4%; Score 546.5; DB 12; Length 348;
Best Local Similarity 42.9%; Pred. No. 5e-32;
Matches 117; Conservative 42; Mismatches 111; Indels 3; Gaps 3;

Qy 34 EDQALTLRLTASDVTKHYELVELGKGYGVKVDLVVYKGTGTQKALKFNKSKTKLKN 93
Db 43 EDMTSLSAQTLVRAEVDLYEVRPLGQGRYGRVLLVTHRQKGTPLAKQLPKPTSLRG 102
Qy 94 FLREVSITNSLSPPFIKVFVDFVVFTEDCYVFAQYAPAGLFDIIPQVGLPDTVKR 153
Db 103 FLYEFVGLSLGSAHSAIVTAAGTGISAHYSFLTEPVLHGDLMAFQPKVGLPQPAVHR 162
Qy 154 CVQOGLALDFMHGROLVHRDIPENLVLFDECRVVKLADFGMTVRVCRVRSVGTIP 213
Db 163 CAAQALSALEYIHAAGLVYRDLPENLVCDPACRRFKLTDFGHTPRGTLRLAGPPI 222
Qy 214 YTAPEVC-QAGRADGLAVDTGVVDVWAFGLIFCVLTGNFPWE-AASGADAFFEFVRWQ 271
Db 223 YTAPELCAPPPLPEGLPIQPALDAWALGVLLFCLLTGYFPWDRPLAEDFPYDFDLTWO 282
Qy 272 GRPL-GLPSQWRRFTEPALRMFORLLALEPERR 303
Db 283 SQQPRDRPQWFGLGAAAADALLGLLDPEHRRR 315

```
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-816-094-4

Query Match
Best Local Similarity 93.1%; Score 2084.5; DB 9; Length 417;
Matches 401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;

QY 1 MSVGCPEPEPPRSRLTCCPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
Db 1 MSVGCPEPEPLHSLFCCGPGAPVPFGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
QY 61 KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSITNSLSSSPFIKVFDDVPET 120
Db 61 KGTGKVDLVAYKGTGTMALKFVNKSKTKLKNFLREVSITNSLSSSPFIKVFDDVPET 120
QY 121 EDCVFAQEAEPAGDLFDIIPQVGLPBDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
Db 121 EECVFAQEAEPAGDLFDIIPQVGLPBDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
QY 181 LFDRECRVRKLADFGMTTRVRCRVKVSIGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG 240
Db 181 LFDRECRVRKLADFGMTTRVRCRVKVSIGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG 240
QY 241 VLIFCVLTGNFPWEAASGADAFFEFVWQGRLPGLPSQWRRTPEALRMFQRLALEP 300
Db 241 VLIFCVLTGNFPWEAASGADAFFEFVWQGRLPGLPSQWRRTPEALRMFQRLALEP 300
QY 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
Db 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
QY 361 SGSRPAPPAGSVPLVPVVPVPPVPPFGLAPQPGPGRTDGRADKSKGQVVLATAI 420
Db 361 SGSRPAPPAGSVPLVPVVPVPPVPPFGLAPQPGPGRTDGRADKSKGQVVLATAI 420
QY 359 SGSRPSPSPVGPV-----VPVPVPVPVPEAGLAPPAPPGRGTGRADKSKGQVVLATAI 413
Db 359 SGSRPSPSPVGPV-----VPVPVPVPVPEAGLAPPAPPGRGTGRADKSKGQVVLATAI 413
QY 421 EICV 424
Db 414 EICV 417

RESULT 8
US-10-233-613-4
; Sequence 4, Application US/10233613
; Publication No. US20030022339A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536DIV
; CURRENT APPLICATION NUMBER: US/10/233,613
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-233-613-4

Query Match
Best Local Similarity 93.1%; Score 2084.5; DB 15; Length 417;
Matches 401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;

QY 1 MSVGCPEPEPPRSRLTCCPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
Db 1 MSVGCPEPEPLHSLFCCGPGAAVPFGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
QY 61 KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSITNSLSSSPFIKVFDDVPET 120
Db 61 KGTGKVDLVAYKGTGTMALKFVNKSKTKLKNFLREVSITNSLSSSPFIKVFDDVPET 120
QY 121 EDCVFAQEAEPAGDLFDIIPQVGLPBDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
Db 121 EECVFAQEAEPAGDLFDIIPQVGLPBDTVKRCVQQLGLALDFMHGRLVHRDIKPNV 180
QY 181 LFDRECRVRKLADFGMTTRVRCRVKVSIGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG 240
Db 181 LFDRECRVRKLADFGMTTRVRCRVKVSIGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG 240
QY 241 VLIFCVLTGNFPWEAASGADAFFEFVWQGRLPGLPSQWRRTPEALRMFQRLALEP 300
Db 241 VLIFCVLTGNFPWEAASGADAFFEFVWQGRLPGLPSQWRRTPEALRMFQRLALEP 300
QY 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
Db 301 ERGPAKEVFRFLKHELTSELRRRSHRARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
QY 361 SGSRPAPPAGSVPLVPVVPVPPVPPFGLAPQPGPGRTDGRADKSKGQVVLATAI 420
Db 361 SGSRPAPPAGSVPLVPVVPVPPVPPFGLAPQPGPGRTDGRADKSKGQVVLATAI 420
QY 359 SGSRPSPSPVGPV-----VPVPVPVPVPEAGLAPPAPPGRGTGRADKSKGQVVLATAI 413
Db 359 SGSRPSPSPVGPV-----VPVPVPVPVPEAGLAPPAPPGRGTGRADKSKGQVVLATAI 413
QY 421 EICV 424
Db 414 EICV 417

RESULT 7
US-09-734-032-4
; Sequence 4, Application US/09734032
; Patent No. US20020103116A1
; GENERAL INFORMATION:
; APPLICANT: WEI et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL00536
; CURRENT APPLICATION NUMBER: US/09/734,032
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60205228
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 417
; TYPE: PRT
; ORGANISM: rat
US-09-734-032-4

Query Match
Best Local Similarity 93.1%; Score 2084.5; DB 10; Length 417;
Matches 401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;
```

```
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THEREOF
; CURRENT APPLICATION NUMBER: US/10/233,613
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-233-613-2

Query Match
Best Local Similarity 100.0%; Score 2239; DB 15; Length 424;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLTAAASDVTKHYELVRELG 60
Db 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLTAAASDVTKHYELVRELG 60
QY 61 KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSIITNSLSSSPFIKVFDDVVFET 120
Db 61 KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSIITNSLSSSPFIKVFDDVVFET 120
QY 121 EDCYVFAOEYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHGROLVHRDIKPNV 180
Db 121 EDCYVFAOEYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHGROLVHRDIKPNV 180
QY 181 LFDRECRVRKVLADFGMTTRVRCRVKRVSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
Db 181 LFDRECRVRKVLADFGMTTRVRCRVKRVSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
QY 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVVPVPEPGLAPQPGRTDGRADKSGQVVLATAI 420
Db 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVVPVPEPGLAPQPGRTDGRADKSGQVVLATAI 420
QY 421 EICV 424
Db 421 EICV 424

RESULT 5
US-10-288-798-1
; Sequence 1, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Dannel B.;
; APPLICANT: WALIA, Narinder K.; HAPALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dying Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Javalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BUTFORD, Neil;
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
```

```
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,495
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 7312543CD1
US-10-288-798-1

Query Match
Best Local Similarity 99.8%; Score 2235; DB 12; Length 424;
Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLTAAASDVTKHYELVRELG 60
Db 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLTAAASDVTKHYELVRELG 60
QY 61 KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSIITNSLSSSPFIKVFDDVVFET 120
Db 61 KGTGKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSIITNSLSSSPFIKVFDDVVFET 120
QY 121 EDCYVFAOEYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHGROLVHRDIKPNV 180
Db 121 EDCYVFAOEYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHGROLVHRDIKPNV 180
QY 181 LFDRECRVRKVLADFGMTTRVRCRVKRVSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
Db 181 LFDRECRVRKVLADFGMTTRVRCRVKRVSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG 240
QY 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVVPVPEPGLAPQPGRTDGRADKSGQVVLATAI 420
Db 241 VLIFCVLGTGNFPWEAASGADAFEEFVWQGRVLPVVPVPEPGLAPQPGRTDGRADKSGQVVLATAI 420
QY 301 ERGPAKEVFRFLKHELTSELRRRPSHRARKPGDRPPAAGPLRLAEPGLKRTVLTESG 360
Db 301 ERGPAKEVFRFLKHELTSELRRRPSHRARKPGDRPPAAGPLRLAEPGLKRTVLTESG 360
QY 361 GSRPAPAVGSPVLPVVPVPEPGLAPQPGRTDGRADKSGQVVLATAI 420
Db 361 GSRPAPAVGSPVLPVVPVPEPGLAPQPGRTDGRADKSGQVVLATAI 420
QY 421 EICV 424
Db 421 EICV 424

RESULT 6
US-09-816-094-4
; Sequence 4, Application US/09816094
; Patent No. US20020064851A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:57:52 ; Search time 38 seconds
(without alignments)
2075.186 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPPRSLTCCPGP.....GRADSKGVVLAIEICV 424

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US02_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2239	100.0	424	9	US-09-816-094-2
2	2239	100.0	424	10	US-09-734-032-2
3	2239	100.0	424	14	US-10-016-985-2
4	2239	100.0	424	15	US-10-233-613-2
5	2235	99.8	424	12	US-10-288-798-1
6	2084.5	93.1	417	9	US-09-816-094-4
7	2084.5	93.1	417	10	US-09-734-032-4
8	2084.5	93.1	417	15	US-10-233-613-4
9	1822	81.4	346	9	US-09-916-790-2
10	546.5	24.4	348	12	US-10-288-798-16
11	546.5	24.4	348	12	US-10-353-690-50
12	394	17.6	256	12	US-09-863-776-40
13	390	17.4	256	11	US-09-823-187-96
14	383.5	17.1	250	12	US-09-976-782-31
15	350	15.6	254	12	US-09-976-782-30

16	350	15.6	256	12	US-09-863-776-41	Sequence 41, Appl
17	349.5	15.6	252	12	US-09-976-782-41	Sequence 41, Appl
18	346	15.5	648	15	US-10-024-036B-5	Sequence 5, Appl
19	346	15.5	817	10	US-09-992-481-4	Sequence 4, Appl
20	346	15.5	817	12	US-10-434-034-4	Sequence 4, Appl
21	345.5	15.4	514	12	US-09-949-029-115	Sequence 115, App
22	345.5	15.4	514	12	US-09-949-029-116	Sequence 116, App
23	341.5	15.3	729	12	US-10-161-565-26	Sequence 26, Appl
24	341.5	15.3	729	15	US-10-142-356-11	Sequence 11, Appl
25	341.5	15.3	729	15	US-10-195-101-33	Sequence 33, Appl
26	341.5	15.3	752	10	US-09-835-081-2	Sequence 27, Appl
27	340.5	15.2	713	12	US-10-161-565-27	Sequence 27, Appl
28	340.5	15.2	744	10	US-09-835-081-4	Sequence 8, Appl
29	340	15.2	513	14	US-10-020-038-8	Sequence 3, Appl
30	337.5	15.1	744	10	US-09-919-585-3	Sequence 28, Appl
31	336.5	15.0	688	12	US-10-161-565-28	Sequence 28, Appl
32	336.5	15.0	688	12	US-10-161-565-29	Sequence 29, Appl
33	336	15.0	765	10	US-09-975-326-4	Sequence 4, Appl
34	336	15.0	765	12	US-10-462-887-4	Sequence 4, Appl
35	336	15.0	765	15	US-10-217-357-4	Sequence 4, Appl
36	336	15.0	795	10	US-09-919-585-12	Sequence 12, Appl
37	336	15.0	795	12	US-10-161-565-25	Sequence 25, Appl
38	336	15.0	795	15	US-10-142-356-9	Sequence 9, Appl
39	335.5	15.0	766	10	US-09-975-326-2	Sequence 2, Appl
40	335.5	15.0	766	10	US-09-934-406-2	Sequence 2, Appl
41	335.5	15.0	766	12	US-10-170-789-21	Sequence 21, Appl
42	335.5	15.0	766	12	US-10-388-798-14	Sequence 14, Appl
43	335.5	15.0	766	12	US-10-462-887-2	Sequence 2, Appl
44	335.5	15.0	766	15	US-10-217-357-2	Sequence 2, Appl
45	335	15.0	460	9	US-09-740-627-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-816-094-2
; Sequence 2, Application US/09816094
; Patent No. US20020064851A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Human
; US-09-816-094-2

Query Match	100.0%	Score	2239;	DB	9;	Length	424;
Best Local Similarity	100.0%	Pred. No.	2.3e-155;				
Matches	424;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MSVGCPEPPRSLTCCPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG	60				
Db	1	MSVGCPEPPRSLTCCPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG	60				
Qy	61	KGTGKVDLVVYKGTGTHWALKFNKSTKLKNEFLREVSITNSLSSSPFIKVFDFVFET	120				
Db	61	KGTGKVDLVVYKGTGTHWALKFNKSTKLKNEFLREVSITNSLSSSPFIKVFDFVFET	120				
Qy	121	EDCVYFAOEYAPAGDLFDIIPQVGLPDDTVKRCVQQLGLALDFMHGRLVHRDIKPNV	180				
Db	121	EDCVYFAOEYAPAGDLFDIIPQVGLPDDTVKRCVQQLGLALDFMHGRLVHRDIKPNV	180				
Qy	181	LLFDRECRVRVKLADFGMTFRYVGRVKRSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG	240				
Db	181	LLFDRECRVRVKLADFGMTFRYVGRVKRSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFG	240				

Search completed: December 5, 2003, 09:58:52
Job time : 23 secs

QY	41	LRTLAASDVTKHYELVRELKGTGYKVDUVTKGTQMAKPVNKS-----TKLKV	93
DB	1	MSTFRQEDVEDHYEKGEEELGSGQFAIVRKCOQKGTMEYAAKFIKKRLLPSSRRRGVSRE	60
QY	94	FUREVSITNSLSSSPFIKVFDFVVFETEDCYVFAEYAPAGOLFDIIPQVGILPDETVR	153
DB	61	PEREVSIIAREI-RHPNITLHP-VFENKTDVILVELVSGGELFDFLAKESLITDEAFO	118

Db 49 SATDEQPHIGNYRLQKTIKGNFAKVLARHVLGTREAVAKIIDKTLQNPISLQKLFREV 108
Qy 99 SITNSLSSPFIKVDVVFETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQOOL 158
Db 109 RIMKIL-NHPNIVKLFVI-ETEKTLVLMVEYASGGEVFDYLVAHGRMKKEAKAFQI 166
Qy 159 GLADFMHGRQLVHRDIKPNVLLPDRCCRVKLADEGMTR--VGCVRKVSCTIPIYTA 216
Db 167 VSAVOYCHQKCIVRHDKAEALLL-DAD-MNKIADFGSNEFTVGNKLDTCGSPPYAA 224
Qy 217 PEVQAGRADGLAVDTGVDVWAFGLVIFCVLTGNFPMWAAAGADAFEEFVRMORGLPG 276
Db 225 PELFGKKYDG---PEVDVWSLVILYTLVSGSLPFDGQNLKE--LRERVLRGKYAVPF 278
Qy 277 LPSQWRRTPEALRMFORLLALEPERRGPAKEVER 311
Db 279 YMS-----TDCENLLKLLVLPNPKRGSLEQIMK 307

RESULT 10
US-08-817-832B-31
; Sequence 31, Application US/08817832B
; Patent No. 6579691
; GENERAL INFORMATION:
; APPLICANT: MANDELKOM, Eckhard, et al.
; TITLE OF INVENTION: No. 6579691el Protein Kinase (NPK-110)
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: US
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,832B
; FILING DATE: 28-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP95/04258
; FILING DATE: 30-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94 11 7122.5
; FILING DATE: 28-OCT-1994
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-817-832B-31

Query Match 14.9%; Score 333.5; DB 4; Length 779;
Best Local Similarity 26.2%; Pred. No. 2.5e-19;
Matches 107; Conservative 72; Mismatches 138; Indels 91; Gaps 16;
Qy 45 AASDVTKH--YELVRELKGTGKVDLVVYKGTCTKMAKLVNKS--TKLKNFLREV 98
Db 35 SATDEQPHIGNYRLQKTIKGNFAKVLARHVLGTREAVAKIIDKTLQNPISLQKLFREV 94
Qy 99 SITNSLSSPFIKVDVVFETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQOOL 158
Db 95 RIMKIL-NHPNIVKLFVI-ETEKTLVLMVEYASGGEVFDYLVAHGRMKKEAKAFQI 152
Qy 159 GLADFMHGRQLVHRDIKPNVLLPDRCCRVKLADEGMTR--VGCVRKVSCTIPIYTA 216
Db 153 VSAVOYCHQKCIVRHDKAEALLL-DAD-MNKIADFGSNEFTVGNKLDTCGSPPYAA 220

Qy 217 PEVQAGRADGLAVDTGVDVWAFGLVIFCVLTGNFPMWAAAGADAFEEFVRMORGLPG 276
Db 211 PELFGKKYDG---PEVDVWSLVILYTLVSGSLPFDGQNLKELRERSCLR-GKYRVPF 265
Qy 277 LPSQWRRTPEALRMFORLLALEPERRGPAKEVER----- 311
Db 266 YMS-----TDCENLLKLLVLPNPKRGSLEQIMKORWMMNVGHEEBELKPYSEPELDND 319
Qy 312 -----FLKHELTSELRRRPSHR-----ARKPP---GDRPPAAGPI----- 343
Db 320 AKRIDIMVTMGFARDEINDALVSKYDEVMATYILLGRKPPPEFEGGESLSSGNLCORSRP 379
Qy 344 -----RLKAPGLK--RTVLTE-----SGSGSRPAPPAVGSVLP 376
Db 380 SSDLNNTLSQPAHLKVQRTISANQKORRFSDHAGPSIPPAVSYTKRP 427

RESULT 11
US-08-924-183-2
; Sequence 2, Application US/08924183A
; Patent No. 6218109
; GENERAL INFORMATION:
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1003
; CURRENT APPLICATION NUMBER: US/08/924,183A
; CURRENT FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-924-183-2

Query Match 14.9%; Score 333; DB 3; Length 476;
Best Local Similarity 28.7%; Pred. No. 1.4e-19;
Matches 102; Conservative 62; Mismatches 138; Indels 54; Gaps 13;
Qy 53 YELVRELKGTGKVDLVVYKGTCTKMAKLVNKSKT--KLKNFLREVSTNSLSSPFI 110
Db 9 WDLVQTLGEGAYGEVQLAVNRITQAVAVKIVDMKRAIDCPQNIKEICI-NKMLSHENV 67
Qy 111 IKVEDVVFETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQOGLALDFMGRQL 170
Db 68 VKFYGHRREGHIQYLFL-EYCSGGELEFDRIEEDIGMEPEQDAQRFHQLMAGVYVYLGIGI 126
Qy 171 VHRDIKPNVLLPDRCCRVKLADEGMTR--RVGCR--VKRVSGTIPYTAPEVCOAGRA 225
Db 127 THRDIKPNVLLPDR--NLKISDFGLATVFRHNNRRELLNKKMGCTLPYVAPELLK--RK 182
Qy 226 DGLAVDTGVDVWAFGLVIFCVLTGNFPMWAAAGADAFEEFVRMORGLPGLPSQWRRT 285
Db 183 EFHA--EPVDVWSGIVLVTAMLAGELFDQPSDS---CQEYSDMKKKTLYLP--WKID 235
Qy 286 EPALRMFORLLALEPERR-----GPAKEVFRFLKHEL 317
Db 236 SAPLALHLKILVETPSARITIPDIKKRWYKNPLNRAKRPATSGMSSSSGFSKH-I 294
Qy 318 TSELRRRP----SHRARKPPGDRPPAAGPLRLAPGLKRTVLTESGSGSRPAPP 368
Db 295 HSNLDFSPVNGSSEETVKFSSSQPEPFTGLSLWDTGFSNVDKLVQGISFSQPTCP 350

RESULT 12
US-09-488-364-2
; Sequence 2, Application US/09488364
; Patent No. 6307015
; GENERAL INFORMATION:
; APPLICANT: Ellledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS


```

; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bosotti, Roberta
; APPLICANT: Scacheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 33
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank Accession No. 6458561 g3089349
US-09-523-849-33

Query Match      15.3%; Score 341.5; DB 4; Length 729;
Best Local Similarity 24.6%; Pred. No. 4.9e-20;
Matches 117; Conservative 75; Mismatches 138; Indels 145; Gaps 18;

QY 37 QALTLRT-----LAASDVTKH---YELVRELKGTGKVDLVVYKGTGTMALK 82
DB 26 QEVTSRSGARCNSASTASCADEPHIGNVLLKTIKGNFAKVKLARHILTGREVAIK 85

QY 83 FVNKSK---TKLKNFLREVSINLSGSPFIKVPDVVFETEDCVFAQEVAPAGDLFDI 139
DB 86 IIDKTQLNPTSLOKLFREVIRMKIL-NHENIVKLFVI-ETEKTYLIMEYASGGEVDY 143

QY 140 IPPQVGLPDTVKRCVQQLGLADFWHGSQLVHRDIKEENVLLFDRECRVKLADFGMTR 199
DB 144 LVAHGRMKKEARSFRQIVSAVQYCHQKRIVHRDLKAENLL-LDAD-MNIKIADFGEN 201

QY 200 R--VCCRKRVSGTTPYTAPEVCOAGRADGLAVDTGVDFVAFGLVFCVLTGNFPWEAAS 257
DB 202 EFTVGKLDTCGSPYAPAPLFOGKYDG----PEVDVMSGLVILYTLVSGSLPFDGON 257

QY 258 GADAFEFVVRQGRGLPGLPSQWRRTPEPALRMFORLLALEPERRGPAKEVFR----- 311
DB 258 LKE--LRERVLAGKYRIPYMS-----TDCENLLKRFVLNPIKRGTLQEIIMKORWINA 309

QY 312 -----FLXHEL-----TS 319
DB 310 GHEDELKPFVPELDDISDQKRIDIMVGMGYSCIEIQESLSKMKYDEITATYLLGRKSS 369

QY 320 EL-----RRRPSHRARKPPGDRP-----PA----- 339
DB 370 ELDASSSSSSNLAKYRPPSSDLNNSTQSPHHKVRQSVSSSQKORRYSDHAGPAIPGV 429

QY 340 -AGPIRL-APGPIKRTVLTE-----SGSGSRPAPPAGVPLPVPPVPV 382
DB 430 VAYPKRSOTSTADGDKEDGISRSKSSGSAVGKGIAPASPMIGNASNPNKADIP 484

RESULT 5
US-08-924-183-8
; Sequence 8, Application US/08924183A
; Patent No. 6218109
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1003
; CURRENT APPLICATION NUMBER: US/08/924,183A
; CURRENT FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-488-364-8

Query Match      15.2%; Score 340; DB 4; Length 513;
Best Local Similarity 32.0%; Pred. No. 4.2e-20;
Matches 85; Conservative 48; Mismatches 115; Indels 18; Gaps 7;

QY 45 AASDVTKHVELVRELKGTGKVDLVVYKGTGTMALKFVN-KSKTKLKNFLREVSITNS 103
DB 14 ATREFVEGWTLAQTLGEGAYGEVKLLINRQTGGCGMKMVDLKKHPDAANSVRKEVCIQ 73

QY 104 LSSSPFIKVPDVVFETEDCVFAQEVAPAGDLFDIIPPQVGLPDTVKRCVQQLGLAD 163
DB 74 MLQDKHILRFFGKRSQGSVEYIFL-EYAAGGELFDRIEFDVGMPPQHEAQRVFTQLLSGLN 132

QY 164 FMHGRQLVHRDIKPEENVLLFDRECRVVLADFGMTRRVGCRVK-----RVSGTIPYTAPE 218
DB 133 YLHQGIAHARDLKPENLLDHD--NVKISDFGMATMFRCKGKELLDKRCGTLIPYVAPE 190

QY 219 VCOAGRADGLAVDTGVDFVAFGLVFCVLTGNFPWEAASGADAFEFVVRQGRGLPGLP 278
DB 191 VLQKAYQPPA-----DLWSCGVILVTMLAGELPDQPS---TNCTEFTNWRDNDHWQLQ 242

QY 279 SQWRRFTPEPALRMFOR-LLALEPERR 303
DB 243 TPWSKLDTLAISLLRKLKLLATSPGTR 268

RESULT 6
US-09-488-364-8
; Sequence 8, Application US/09488364
; Patent No. 6307015
; GENERAL INFORMATION:
; APPLICANT: Elledge, Stephen J.
; APPLICANT: Sanchez, Yolanda
; TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
; FILE REFERENCE: 120541-1013
; CURRENT APPLICATION NUMBER: US/09/488,364
; CURRENT FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-488-364-8

Query Match      15.2%; Score 340; DB 4; Length 513;
Best Local Similarity 32.0%; Pred. No. 4.2e-20;
Matches 85; Conservative 48; Mismatches 115; Indels 18; Gaps 7;

QY 45 AASDVTKHVELVRELKGTGKVDLVVYKGTGTMALKFVN-KSKTKLKNFLREVSITNS 103
DB 14 ATREFVEGWTLAQTLGEGAYGEVKLLINRQTGGCGMKMVDLKKHPDAANSVRKEVCIQ 73

QY 104 LSSSPFIKVPDVVFETEDCVFAQEVAPAGDLFDIIPPQVGLPDTVKRCVQQLGLAD 163
DB 74 MLQDKHILRFFGKRSQGSVEYIFL-EYAAGGELFDRIEFDVGMPPQHEAQRVFTQLLSGLN 132

QY 164 FMHGRQLVHRDIKPEENVLLFDRECRVVLADFGMTRRVGCRVK-----RVSGTIPYTAPE 218
DB 133 YLHQGIAHARDLKPENLLDHD--NVKISDFGMATMFRCKGKELLDKRCGTLIPYVAPE 190

QY 219 VCOAGRADGLAVDTGVDFVAFGLVFCVLTGNFPWEAASGADAFEFVVRQGRGLPGLP 278
DB 191 VLQKAYQPPA-----DLWSCGVILVTMLAGELPDQPS---TNCTEFTNWRDNDHWQLQ 242

QY 279 SQWRRFTPEPALRMFOR-LLALEPERR 303
DB 243 TPWSKLDTLAISLLRKLKLLATSPGTR 268

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:55:26 ; Search time 21 Seconds
(without alignments)
854.276 Million cell updates/sec

Title: US-10-016-985-2
Perfect score: 2239
Sequence: 1 MSVGCPEPPPSRLTCCGPG.....GRADKSGQVVLATAFICV 424

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	100.0	424	4	US-09-816-094-2
2	2084.5	93.1	417	4	US-09-816-094-4
3	341.5	15.3	729	2	US-08-677-298-2
4	341.5	15.3	729	4	US-09-523-849-33
5	340	15.2	513	3	US-08-924-183-8
6	340	15.2	513	4	US-09-488-364-8
7	336	15.0	765	4	US-09-975-326-4
8	335.5	15.0	766	4	US-09-975-326-2
9	334	14.9	793	4	US-09-523-849-32
10	333.5	14.9	779	3	US-08-817-832B-31
11	333	14.9	476	3	US-08-924-183-2
12	333	14.9	476	4	US-09-488-364-2
13	328.5	14.7	448	2	US-09-159-385-2
14	328.5	14.7	448	3	US-09-186-277-2
15	328	14.6	476	3	US-09-522-800-16
16	328	14.6	476	3	US-08-924-183-1
17	328	14.6	476	4	US-09-488-364-1
18	328	14.6	476	4	US-09-113-785-1
19	323.5	14.4	722	4	US-09-984-890-4
20	323.5	14.4	722	4	US-08-817-832B-32
21	322.5	14.4	483	3	US-08-924-183-9
22	322.5	14.4	483	4	US-09-488-364-9
23	322.5	14.4	745	4	US-09-523-849-36
24	321.5	14.4	724	4	US-09-984-890-2
25	318	14.2	454	2	US-09-159-385-1
26	318	14.2	454	3	US-09-186-277-1
27	315	14.1	552	3	US-08-557-006C-40

28	311.5	13.9	582	4	US-09-428-711A-2	Sequence 2, Appli
29	311.5	13.9	1051	4	US-09-428-711A-14	Sequence 14, Appli
30	310	13.8	668	4	US-09-930-181-2	Sequence 2, Appli
31	309	13.8	633	3	US-08-557-006C-43	Sequence 43, Appli
32	308.5	13.8	257	2	US-07-857-224B-25	Sequence 25, Appli
33	307	13.7	543	4	US-09-529-093A-2	Sequence 2, Appli
34	307	13.7	543	4	US-09-529-154-2	Sequence 2, Appli
35	305.5	13.6	476	3	US-08-924-183-7	Sequence 7, Appli
36	305.5	13.6	476	4	US-09-488-364-7	Sequence 11, Appli
37	302	13.5	631	4	US-09-523-849B-11	Sequence 34, Appli
38	302	13.5	776	4	US-09-523-849-34	Sequence 1, Appli
39	298.5	13.3	604	3	US-09-101-146-1	Sequence 35, Appli
40	298	13.3	604	4	US-09-523-849-35	Sequence 6, Appli
41	295.5	13.2	257	3	US-09-101-146-6	Sequence 2, Appli
42	294.5	13.2	396	4	US-09-801-876B-2	Sequence 16, Appli
43	294.5	13.2	1050	4	US-09-428-711A-16	Sequence 2, Appli
44	294	13.1	436	4	US-09-734-673-2	Sequence 2, Appli
45	294	13.1	436	4	US-09-523-849-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-816-094-2
; Sequence 2, Application US/09816094
; Patent No. 6534299
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Human
US-09-816-094-2

Query Match 100.0%; Score 2239; DB 4; Length 424;
Best Local Similarity 100.0%; Pred. No. 5.4e-177; Indels 0; Gaps 0;
Matches 424; Conservative 0; Mismatches 0;

QY	1	MSVGCPEPPPSRLTCCGPGTAPGPGAGVPLLTEDMQALTLETLAASDVTKHYELVRELG	60
DB	1	MSVGCPEPPPSRLTCCGPGTAPGPGAGVPLLTEDMQALTLETLAASDVTKHYELVRELG	60
QY	61	KGTQKVDLVVYKGTGTQKALFKVNSKTKLKNFLREVSITNSLSSPFIKVFVDVVFET	120
DB	61	KGTQKVDLVVYKGTGTQKALFKVNSKTKLKNFLREVSITNSLSSPFIKVFVDVVFET	120
QY	121	EDCYFAQYAPAGDLFDIIPQVGLPEDEVKRCVQQLGLALDFMGRQLVHRDIKENV	180
DB	121	EDCYFAQYAPAGDLFDIIPQVGLPEDEVKRCVQQLGLALDFMGRQLVHRDIKENV	180
QY	181	LLFDRECRVVKLADFGMTRRVGCRVVKRVSGTIPYTAPEVCQAGRADGLAVDTGVWVAFG	240
DB	181	LLFDRECRVVKLADFGMTRRVGCRVVKRVSGTIPYTAPEVCQAGRADGLAVDTGVWVAFG	240
QY	241	VLIFCVLTGNFWEAASGADAFEEFVRVQGRGLPGLPSQWRFRFTFPAIRMFORLLALBP	300
DB	241	VLIFCVLTGNFWEAASGADAFEEFVRVQGRGLPGLPSQWRFRFTFPAIRMFORLLALBP	300
QY	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPAAAGPLRLLEAPGLKRTVLITESG	360
DB	301	ERRGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPAAAGPLRLLEAPGLKRTVLITESG	360
QY	361	SGSRPAPPVAVGSVPLPVFPVPVPEPGIAPQPPGRTDGRADSKSGQVVLATAI	420

5/17/09
60205228

Db 1386 GCGAGGACCGAGCGCGTGCAGACAGAGCAAGGCGAGGTGGTATTGGCCACAGCCATC 1445
QY 1261 GAGATCTGCGTCTGA 1275
Db 1446 GAGATCTGCGTCTGA 1460

RESULT 15
AX589475
LOCUS AX589475 936 bp DNA linear PAT 24-JAN-2003
DEFINITION Sequence 4 from Patent WO02083882.
ACCESSION AX589475
VERSION AX589475.1 GI:27900911
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Koehler, R.H.
Regulation of human serine/threonine protein kinase-like protein
Patent: WO 02083882-A 4 24-OCT-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
1..936
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 155 a 308 c 308 g 165 t
ORIGIN

Query Match 73.0%; Score 931.2; DB 6; Length 936;
Best Local Similarity 99.7%; Pred. No. 2.7e-109;
Matches 933; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGAGCGTGGGTGGCCAGAGCGTGAAGCGCGCGCGTCCCTGACCTGCTGTGGCGCGGGG 60
Db 1 ATGAGCGTGGGTGGCCAGAGCGTGAAGCGCGCGCGTCCCTGACCTGCTGTGGCGCGGGG 60
QY 61 ACTGCCCTGGGCGCTGGTGGCGGTGGCCCTTCTCACTGAAGACATGACAGCCCTGACT 120
Db 61 ACTGCCCTGGGCGCTGGTGGCGGTGGCCCTTCTCACTGAAGACATGACAGCCCTGACT 120
QY 121 CTCGCGACACTGGCGCCAGCGACGCTACCAAGCACTAGCAACTAGTCCGGAGCTGGGC 180
Db 121 CTCGCGACACTGGCGCCAGCGACGCTACCAAGCACTAGCAACTAGTCCGGAGCTGGGC 180
QY 181 AAAGGACCTATGGAGAGTTGACCTGGTGGTCTACAAGGGCACAGGCACAAATGGCA 240
Db 181 AAAGGACCTATGGAGAGTTGACCTGGTGGTCTACAAGGGCACAGGCACAAATGGCA 240
QY 241 CTGAAGTTTGTGAACAAGCAAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATC 300
Db 241 CTGAAGTTTGTGAACAAGCAAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATC 300
QY 301 ACCAACAGCCTCTCCAGCCCTTTCATCAAGGTCTTTGACGTGTCTTTGAGACA 360
Db 301 ACCAACAGCCTCTCCAGCCCTTTCATCAAGGTCTTTGACGTGTCTTTGAGACA 360
QY 361 GAGGACTGTACGTCTTTGGCCAGGAGTAGGACCTGCTGGGAGCCTGTTGACATATC 420
Db 361 GAGGACTGTACGTCTTTGGCCAGGAGTAGGACCTGCTGGGAGCCTGTTGACATATC 420
QY 421 CTTCCCCAGGTGGGGTCCCTGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCCTG 480
Db 421 CTTCCCCAGGTGGGGTCCCTGAGGACACGGTGAAGCGTGTGTGACAGAGCTGGGCCTG 480
QY 481 GCGCTGGACTTTCATGACGGCGGCGAGCTGGTGCACCGGACATCAAGCCCGAGAACTG 540
Db 481 GCGCTGGACTTTCATGACGGCGGCGAGCTGGTGCACCGGACATCAAGCCCGAGAACTG 540
QY 541 CTCTGTTCGACCGGAGTGGCGCGGTAAAGCTGGCGGACTTTCGGCATGACCGCGCGC 600

Db 541 CTCTGTTCGACCGGAGTGGCGCGGTAAAGCTGGCGGACTTTCGGCATGACCGCGCGC 600
QY 601 GTGGGCTGCGCGCGTCAAGCGCGTGAAGCGGCAACCAATCCCTTACACGGCGCCTGAGGTGTC 660
Db 601 GTGGGCTGCGCGCGTCAAGCGCGTGAAGCGGCAACCAATCCCTTACACGGCGCCTGAGGTGTC 660
QY 661 CAGCGGCGCGCGCGGAGCGGCGTGGACACGGGCGGTGGACGCTGTGGGCGCTTCGGC 720
Db 661 CAGCGGCGCGCGCGGAGCGGCGTGGACACGGGCGGTGGACGCTGTGGGCGCTTCGGC 720
QY 721 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGTTGGGAGGCGGCTTCGGGCGCCGAC 780
Db 721 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGTTGGGAGGCGGCTTCGGGCGCCGAC 780
QY 781 GCCTTCTTCGAGGAGTTGCTGCTGCTGAGCGGCGCGCTTCGGGCGCTTCGCTTCGAG 840
Db 781 GCCTTCTTCGAGGAGTTGCTGCTGCTGAGCGGCGCGCTTCGGGCGCTTCGCTTCGAG 840
QY 841 TGGCGCGCTTTCACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGGCGCTGAGGCC 900
Db 841 TGGCGCGCTTTCACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGGCGCTGAGGCC 900
QY 901 GAGCGCGCGCGCGCCAGCAAGGAGGTGTTCCGCTTC 936
Db 901 GAGCGCGCGCGCGCCAGCAAGGAGGTGTTCCGCTTC 936

Search completed: December 7, 2003, 16:12:00
Job time : 3236 secs

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE

1 (sites)
Rattus.
Nara,K.
Cloning and characterization of a novel serine/threonine protein
kinase gene expressed dominantly in developing brain
Published Only in Database (1999)
JOURNAL
2 (bases 1 to 1529)
Nara,K.
Direct Submission
Submitted (07-JAN-1998) Kiyomitsu Nara, Mitsubishi Kasei Institute
of Life Sciences, Department of Glycobiology; MinamiOoya 11,
Machida, Tokyo 194, Japan (E-mail:nara@libra.lis.m-kagaku.co.jp,
Tel:81-427-24-6298, Fax:81-427-24-6316)
JOURNAL

FEATURES

Location/Qualifiers
1..1529
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="hippocampus"
/dev_stage="postnatal 7-day"
1..1529
/gene="sbk"
207..1460
/gene="sbk"
/EC_number="2.7.1.-"
/note="old gene name:PKN"
/codon_start=1
/product="serine/threonine protein kinase"
/protein_id="BAA36362.1"
/db_xref="GI:4115429"
/translation="MSVGGPEPELHSLPCGPGAAVPVGGVPLLTEDMOALRLTL
AASDVTKEVLEKGTGYGVLDVAVKGTGKWLKFNKSKTKLKNFLREVSEITN
SLSSPFPIKVDVYFEECYFAQEVAPAGDLFIIPPOVGLPEDTVKVCQQLGL
ALDFHRSQQLVHDIPENVLDFDRECRKVLADGMRVRGCRVKRVSGLIPYTAPE
VQAGRADGFVDTGVWVAFGLIFCVLTGNFPAEASGADAFEEFVRWGRGLPG
LPSQWRFTPEPALLFORLALPEPRGPAKEVFRFLKHLTSELRSPSHRARPFG
DLPLGLRLAEPGLKRLTVLTSGSGSRPSPSVGVPVVPVFPVPEAGLAPAP
PGRTDGRADKSKQGVLAFAIEICV"
363..1157
/gene="sbk"
/note="protein_kinase domain"
384..407
/gene="sbk"
/note="NP_binding domain"
450..452
/gene="sbk"
/note="ATP_binding domain"
726..728
/gene="sbk"
/note="active_site"
1293..1388
/gene="sbk"
/note="Proline-rich region"
1371..1385
/gene="sbk"
/note="SH3-binding core"
BASE COUNT 263 a 495 c 503 g 268 t
ORIGIN

Query Match 75.6%; Score 963.4; DB 10; Length 1529;

Best Local Similarity 86.1%; Pred. No. 2e-113;
Matches 1098; Conservative 0; Mismatches 156; Indels 21; Gaps 2;

QY 1 ATGAGCGTGGCTCCAGAGCCTGAGCCGCCCGCTCCCTGACTGCTGNGSCGCGGG 60

DB 207 ATGAGCGTGGCTCCCTGAGCCTGAAACCGCTCCACTCCCTGCTGCTGGGCGCGGG 266

QY 61 ACTCCCTGGCTGGTGGCGGTGGTGGCCCTTCTCACTGAAGACATGCAAGCCCTGACT 120

DB 267 GCCGCCCTGTACTGGTGGAGGTGTGCCCTCTCTACAGNAGACATGCAAGCCCTGACC 326

QY 121 CTCCGCACTGCGCGCAGCGAGCTGCAACAAGCACTACGAACCTAGTCCGGAGCTGGGC 180
DB 327 CTCCGCACTGCGCTGCCAGCGACGTCACAAGCACTACGAGCTCGTGGGGAGCTGGGT 386
QY 181 AAAGGCACTATCGGAAGGTTGACCTGGTGTCTACAAGGGCACAGGCACAAATAATGGCA 240
DB 387 AAAGGCACTACGGGAAGGTCGACCTGGTGTCTACAAGGGCACAGGCACAAATAATGGCC 446
QY 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTCTACGGGAGGTGAGCATC 300
DB 447 CTGAATTTGTGAATAAGAGTAAGACCAAGCTGAAGAACTTCTCTGCTGAGGTGAGCATC 506
QY 301 ACCAAGACCTCTCTCCAGCCCTTCATCATCAAGTCTTTGACGCTGTTTGTAGACA 360
DB 507 ACCAAGACCTGCTGCTAGCCCTTCATCATCAAGTCTTTGACGCTGTTTGTAGACA 566
QY 361 GAGGACTGCTACGCTTTTGGCCAGGAGTAGGCACCTCTCTGGGGACCTGTTTGTACATCATC 420
DB 567 GAGGAGTGTCTATGCTTTTGTCTAGGAGTATGCACCTCTCTGGGACCTGTTTGTACATCATC 626
QY 421 CTTCCCAAGTGGGCTCCCTGTAGGACAAGGTGAAGCGCTGTGTGACGACGCTGGGCTG 480
DB 627 CTTCTCAGGTGGGCTCCCGGAGGACACGCTGAAGCGCTGTGTGACGACGCTGGGCTG 686
QY 481 GCGGTGACCTTCATGCAAGCGCGGCGAGCTGTGTGACCGGACATCAAGCCCGAGAACTG 540
DB 687 GCACTGGACTTCATGCAATAGCAGGACGCTGTGTGACCGGACATCAAGCCCGAGANTGTG 746
QY 541 CTGCTGTTTCGACACCGGAGTCCCGCGGTAAAGTGGCGGACTTTGGGATGACCGCGCGC 600
DB 747 CTGCTGTTTACCGTGTAGTCCCGCGGTGAAGTGGCTGACTTCGGCATCACCGCGCGC 806
QY 601 GTGGCTCCCGCTCAAGCGGCTGAGCGGACCATCTCTACAGCGGCTGAGTGTGTC 660
DB 807 GTAGGCTCCGTGTGAAGCGAGTAGCGGCACTATACCTACAGCGGCTGAGTGTGTC 866
QY 661 CAGCGGCGCGCGCGCGGCTGCGGTGGACACCGGCGCTGGAGCTGTGGGCTTCGCGC 720
DB 867 CAGGCTGCGCGCGCGCGGCTTCCGCTGTGACACCGGCGCTGGATGTGTGGGCACTTCGCG 926
QY 721 GTGCTCATCTTCTGCGTGTCTCAACCGGCACTTCCCGTGGGAGGCGGCTGTGGGCGCGCAC 780
DB 927 GTGCTCATCTTCTGCGTGTCTCACTGGCACTTCCCGTGGGAGGCTGCTGAGTGGCGAT 986
QY 781 GCCTTCTTCGAGGAGTGTGTGGCTGCGAGCGGCGCGCTGCGCGGCTGCTGCTCGAG 840
DB 987 GCCTTCTTCGAGGAAATTTGTGCGTGGAGGCGGCTGCGCGGCTGCGCATCCAG 1046
QY 841 TGGCGCGCTTCAACCGAGCCCGCTGCGCATGTTCCAGCGCTTACTGGCCCTGGAGGCC 900
DB 1047 TGGCGAGCTTTTACGAGGCTGCTCTACGATGTTCCAGCGGCTTCTGGCGCTGGAGGCT 1106
QY 901 GAGCGCGCGCGCGCGAGCCAAAGAGGTGTTCCGCTTCTTCAAGCACAGTCACTCATCTGAG 960
DB 1107 GAGCGCGTGGGCGCGCGCAAGGAGGTCTTTTCGCTTCTTCAAGCATGAGTCACTCATCTGAG 1166
QY 961 CTGGCGCGCGCGCTCGCACCGCGCGCAAGCCCGCGGAGCGCGCGCGCGCGCGCGCG 1020
DB 1167 CTGGCG 1220
QY 1021 GGGCCACTCGGCTCGAGGCGCTGGGCGGCTCAAGCGGACGCGTGTGACCGAGAGCGGC 1080
DB 1221 GGGCCCTTGGGCTTGAAGGCTCCAGGGCTCACTCAAGCGACATGTGCTCACCGAGAGTGGC 1280
QY 1081 AGCGGCTCCCG 1140
DB 1281 AGCGGCTCGGCGCTTCCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1325
QY 1141 GTGCCAGTCCCGTGGCGGTGCTGTGCGGAGCGCGGCTGCTCCCGAGGGCGCGCGCGCG 1200
DB 1326 GTGCCAGTGGCGGTGCGAGTACCGGTAAGGCTGGTGTGCTCCACCGAGCGCGCGCG 1385
QY 1201 GCGCGGACCGAGCG 1260

		/db xref="GI:19388002"	
		/db xref="LocusID:104175"	
		/translaction="MSVGCPEPEPLSLPCCGGAAPVPGAGVPLLTEDNQAALTRL	
		AASDVKHYELVRELKGTGKVDLVAYGTGKTKMALKFKVNGSKTKLKNFLREVSITN	
		SLSSSPFIKVFVDVPEECYVFAQYAPAGLFDIIPQVGLPEDTKVRCVQOGL	
		ALDFMSRQVJHVDIIPENVILLFDRECRVYKLDLFGMTRVAGCRVKSVTIPTYAPE	
		VCQAGRADGFADVDVDFVAFGLVILFCVLITGNFWEASADAFEFVFNQGRPLG	
		LPSQWRFTPALRMFORLALPEFRGPAKEVFRFLKHELTSELRRRPPHARKPPG	
		DLPLGSLRPAQPLKATVLTESGSRSPSPVGPVVPVFPVFPVPEAGLAPP	
		PGRTDGTDSKQGVNLATAIEICV"	
BASE COUNT		810 a 1193 c 1145 g 805 t	
ORIGIN			
Query Match		75.8%; Score 966.6; DB 10; Length 3953;	
Best Local Similarity		86.3%; Pred. No. 6.1e-114;	
Matches 1100; Conservative		0; Mismatches 154; Indels 21; Gaps 2;	
Qy	1	ATGAGCGTGGCTGCCAGACCTGAGCCGCGCCGCTCCCTGACCTGCTGTGGGCGGGG	60
Db	326	ATGAGCGTGGCTGCCAGACCTGAAACCGCTCCACTCCCTGCTGTGGGCGGGG	385
Qy	61	ACTGCCCTGGCTGTGGCGGTGTGCCCTTCTCACTGAAGACATGCAAGCCCTGACT	120
Db	386	GCGGCCCTGTACAGGTGCAAGTGTGCCCTCTCTCAGAAACATGCAAGCGCTGACC	445
Qy	121	CTCCGCACTGGCGCGCAGCAGCTCACCAAGCACTACAACTAGTCCGGGAGCTGGC	180
Db	446	CTACGCACTGGCTGCCAGCGATGTTACCAAGCACTATGAGCTGTGCGGGAGCTGGC	505
Qy	181	AAAGGCACTATGGAAGTTGACTGTGTGTCTCAAGGGCCACAGGCACAAAATGCA	240
Db	506	AAAGGCACTACGGAAGCTGCACTGTGTGTCTCAAGGGCCACAGGCACAAAATGCC	565
Qy	241	CTGAAGTTGTGAACAAGCAAAACCAAGCTGAAAGAACTTCTTACGGGAGGTGAGCATC	300
Db	566	CTGAATTTGTGAATAAGATGAGCAAAAGCTGAAAGAACTTCTGCGTGAAGTGAATC	625
Qy	301	ACCAACAGCTCTCTCCAGCCCTTCAATCAATCAAGTCTTTGAGCTGTCTTTGAGACA	360
Db	626	ACCAACAGCTCTCTGCTAGCCCTTCAATCAATCAAGTCTTTGAGCTGTCTTTGAGACC	685
Qy	361	GAGGCTGCTAGCTTTTGGCCAGAGTACGACCTGCTGGGAGCTTTTGACATCATC	420
Db	686	GAGGAGTGTATGTCTTTTGGCCAGAGTATGACCTGCTGGGAGCTTTTGACATCATC	745
Qy	421	CTCTCCAGGTGGGCTCTCTGAGCACAGGTGAAGCGCTGTGTGAGCAGCTGGGCGCTG	480
Db	746	CTCTCTCAGGTGGGCTCTCCAGAGCACAGGTGAAGCGCTGTGTGAGCAGCTGGGCGTG	805
Qy	481	GCGCTGGACTTCATGACAGGGCGGAGCTGGTGTGCAACCGGACATCAAGCCGAGAACGTG	540
Db	806	GCGCTGGACTTCATGACAGCAGGAGCTGGTGTGCAACCGGACATCAAGCCGAGAACGTG	865
Qy	541	CTGCTGTTGACCGGAGTGGCGCGCTTAAGCTGGCGACTTGGCATGACGGCGCGC	600
Db	866	CTGCTGTTGACCGGAGTGGCGCGCTGAAAGCTGGCTGACCTGGCATGACGGCGCGC	925
Qy	601	GTGGGCTGGCGCTCAAGCGCTGAGCGGCACCATCCCTTACACGGCGCTGAGGTGTGC	660
Db	926	GTGGGCTGGCGCTGAGCGCTGAGCGGCACATATACCTACACAGCAGCGGTGTGC	985
Qy	661	CAGGGCGGCGCGCAGCGGCTGGCGGTGGACACGGGCTGACCTGTGGGCTTGGC	720
Db	986	CAGGCTGGCGCGCGATGGCTTGGCGGTGGACACGGGCTGACCTGTGGGCTTGGC	1045
Qy	721	GTGCTCATCTTCTGCTGCTACCGCAACTTCCCTGGAGGCGGCTGCGGCGCGAC	780
Db	1046	GTGCTCATCTTCTGCTGCTACCTGCAACTTCCCTGGAGGCTGATCGGGTGGGAT	1105
Qy	781	GCCTTCTTGAGAGTGTGTGCTGGCAGCGGGCGCGCTGCGGGGCTGCTTGGCAG	840
Db	1106	GCCTTCTTGAGAGTGTGTGCTGGCAGCGGGCTGCGCTGCGGGGCTGCAATCGCAG	1165

Qy	841	TGGCGCGCTTACCGAGCCCGCGCTCGCATTTCCAGCGCTTACTGGCCCTGGAGCCC	900
Db	1166	TGGCGCGCTTACCGAGCCCGCGCTCGCATTTCCAGCGCTTACTGGCCCTGGAGCCT	1225
Qy	901	GAGCGCGCGCGCCAGCAAGAGAGGTTCCTTCTCAAGCACGAGCTCACTGCGAG	960
Db	1226	GAGCGCGGTGGGCGCGCCAGAGAGGTCTTCTCGCTTCTCAAGCACGAGCTCACTGCGAG	1285
Qy	961	CTGCGCGCGCGCGCTCGACCCCGCGCGCAAGCCCGCGGAGCCCGCGCGCGCGCC	1020
Db	1286	CTGCGCGCGCGCGCTTCCCAACCGTGCAAGCGCGCTGGGACCGCGCTGC-----CA	1339
Qy	1021	GGCGCATGCGCTCGAGGCGCTGGGCGCTCAAGCGGAGGTGCTGACCGAGAGCGGC	1080
Db	1340	GGGTCCCTGCGCTTGAAGTCCCGGCCACTCAAGCGGACTGTGCTCACCGAGAGTGGC	1399
Qy	1081	AGCGGCTTCCCGCGCGCGCGCGCGCTGCGGTGCGGTGCGCTTGGCGCGGTGCGG	1140
Db	1400	AGCGGCTGCGCGCTTCCCGCACCGCTAGGGCGCGT-----GGTACCC	1444
Qy	1141	GTGCCAGTGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG	1200
Db	1445	GTGCCAGTGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG	1504
Qy	1201	GGCGGACCGACGCGCGCGCGCGCAAGAGCAAGGAGGAGGTGCTGCGGACCGCGCATC	1260
Db	1505	GGCAGGACCGACGCGCGCGCGCAAGAGCAAGGAGGAGGTGCTGCGGACCGCGCATC	1564
Qy	1261	GAGATCTGCGCTTGA	1275
Db	1565	GAGATCTGCGCTTGA	1579
RESULT 12			
BC031759			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			

BC031759 4123 bp mRNA linear ROD 16-APR-2003
Mus musculus SH3-binding kinase, mRNA (cdna clone MGC:28537
IMAGE:4194152), complete cds.
BC031759
BC031759.1 GI:21594543
MGC.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4123)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bortolotto, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Udutin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Heaton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalek, U., Smal, M.A.,
Schnurch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 4123)
Strausberg, R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian


```
|||||
650 CCTCTCAGTGGGGCTCCAGAGGACACGGTGAAGCGCTGTGTGCAGCAGCTGGGGCTG 709
QY
481 GCGCTGACATTCATGACGCGGCGCAGCTGTGTGCACCGACATCAAGCCGAGACGCTG 540
Db
710 GCGCTGACATTCATGACGCGGCGCAGCTGTGTGCACCGACATCAAGCCGAGATGTG 769
QY
541 CTGCTGTTCCAGCGCGAGTGGCGCGGTAAAGCTGGCCGACTTTCGGCATGAGCGCGCC 600
Db
770 CTGCTGTTTGAACGCTGAGTGGCGCGGTGTGAAGCTGCTGCTCGCATGAGCGCGCC 829
QY
601 GTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACCAATCCCTTACACGGCGCCTAGGTGTC 660
Db
830 GTGGGCTGCGCGTGAAGCGCGTGAAGCGGCACTATACGATACACAGCACCGAGGTGTC 889
QY
661 CAGCGGGCGCGCGCGCTGCGGTGACACGCGGCTGACGCGGTGAGCTGTGGGCTTCGCG 720
Db
890 CAGCTGGCGCGCGCGATGGCTTCGCGTGAACAGCGGCGTGAAGCTGTGGGCTTCGCG 949
QY
721 GTGCTCATCTTCTCGTCTCACCGGCAACTTCCCGTGGGAGGCGGCTCGGGCGCCGAC 780
Db
950 GTGCTCATCTTCTGTGTCTCACTGGCACTTCCCGTGGGAGGCTGCATCGGTGGCAT 1009
QY
781 GCCTTCTTCGAGGATTCGTGCGCTGGCAGCGGGGCGCCCTGCGGGGCTGCTTCGCGAG 840
Db
1010 GCCTTCTTCGAGGATTCGTGCGCTGGCAGCGGGGCTGCGCTGCGGGGCTGCCATCGCAG 1069
QY
841 TGGGCGCGCTTCAACGAGCGCGCGTGCATGTTCCAGCGCTTACTGGCCCTGAGAGCC 900
Db
1070 TGGGCGCGCTTACCGAGCGCTGCGCTGCGTATGTTCCAGCGGCTTCTGGCGTAGAGCT 1129
QY
901 GAGCGCGCGCGCCAGCAAGGAGGTGTTCCGCTTCTCAAGCAGAGCTCAGTCCGAG 960
Db
1130 GAGCGCGGTGGCGCGCCGCAAGGAGGTCTTTCGCTTCTCAAGCAGAGCTCAGTCCGAG 1189
QY
961 CTGCGCGCGCGCGCTTCCAGCGCGGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCC 1020
Db
1190 CTGCGCGCGCGCGCTTCCAGCGCGGCTTCCAGCGCGCGCGCGCGCGCGCGCGCGCG 1243
QY
1021 GGGCACTGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGAGCGTGTGACCGAGAGCGCG 1080
Db
1244 GGGTCCCTGCGCTTGAAGCTCCCGGCGCACATCAAGCGCACTGTCTACCGAGAGTGGC 1303
QY
1081 AGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db
1304 AGCGGCTCCCGCGCGCTTCCAGCGCGGCTTCCAGCGCGCGCGCGCGCGCGCGCGCG 1348
QY
1141 GTGCCAGTCCCGTGGCGGCTGTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Db
1349 GTGCCAGTCCCGTGGCGGCTGTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1408
QY
1201 GCGCGGACCGAGCGCGCGCGCGGACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Db
1409 GCGAGACCGAGCGCGCGCGCGGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1468
QY
1261 GAGATCTGGCTTGA 1275
Db
1469 GAGATCTGGCTTGA 1483

RESULT 11
BC025837
LOCUS BC025837
DEFINITION Mus musculus SH3-binding kinase, mRNA (cdna clone MGC:37964
IMAGE:5135723), complete cds.
VERSION BC025837.1 GI:19388001
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3953)
```

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wernley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Soergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickerson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, N.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3953)

Strausberg, R.

Direct Submission

Submitted (11-MAR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbe-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

cDNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nee, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 57 Row: n Column: 20

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

FEATURES

Location/Qualifiers

1..3953

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:37964 IMAGE:5135723"

/tissue_type="Liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_L19"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..3953

/gene="Sbk-pending"

/db_xref="LocusID:104175"

/db_xref="MGI:2135937"

326..1579

/codon_start=1

/product="SH3-binding kinase"

/protein_id="AAH25837.1"

gene

CDS

QY	425	CCAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACAGAGCTGGGCTGGCGC	484
Db	464	CCAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACAGAGCTGGGCTGGCGC	523
QY	485	TGACTTTATGACAGGGCGGCGAGTGTGTGACCGGACATCAAGCCCGAGAAAGTGTGC	544
Db	524	TGGACTTCATGACAGGGCGGCGAGTGTGTGACCGGACATCAAGCCCGAGAAAGTGTGC	583
QY	545	TGTTCGACCGGAGTGCAGCGGCTGAAGCTGGCCGACTTTCGGCATGACGCGCGGTGG	604
Db	584	TGTTGACCGGAGTGCAGCGGCTGAAGCTGGCCGACTTTCGGCATGACGCGCGGTGG	643
QY	605	GCTGCCCGCTCAAGCGCTGAGCGGACCAATCCCTTACAGGGCGCTTGAGGTGTGCCAGG	664
Db	644	GCTGCCCGCTCAAGCGCTGAGCGGACCAATCCCTTACAGGGCGCTTGAGGTGTGCCAGG	703
QY	665	CGGCGCGGCGAGCGGCTGGCGTGACACGGGCTGGAGCTGTGGGCTTCGGCGTGC	724
Db	704	CGGCGCGGCGAGCGGCTGGCGTGACACGGGCTGGAGCTGTGGGCTTCGGCGTGC	763
QY	725	TCATCTTCTGCTGCTCACCGGCAACTTCGCTGGGAGGGCGCTCGGGGCGCGACGCT	784
Db	764	TCATCTTCTGCTGCTCACCGGCAACTTCGCTGGGAGGGCGCTCGGGGCGCGACGCT	823
QY	785	TCCTTCGAGGAGTTCGTGCGCTGGCAGCGGGCGCCCTGCGGGGCTGCTTCGCAGTGGC	844
Db	824	TCCTTCGAGGAGTTCGTGCGCTGGCAGCGGGCGCCCTGCGGGGCTGCTTCGCAGTGGC	883
QY	845	GCGGCTTACGAGCGCGCTGGCGATGTTCCAGCGCTTACTGGCGCTGGAGCCCGAGC	904
Db	884	GCGGCTTACGAGCGCGCTGGCGATGTTCCAGCGCTTACTGGCGCTGGAGCCCGAGC	943
QY	905	GCCGCGGCGAGCCAAAGAGGTGTCGCTTCCCTCAAGCAGAGCTCACGTCGAGCTGC	964
Db	944	GCCGCGGCGAGCCAAAGAGGTGTCGCTTCCCTCAAGCAGAGCTCACGTCGAGCTGC	1003
QY	965	GCCGCGGCGCTCGACCGCGCGCAAGCGCCCGCGGGGACCGCGCGCGCGCGCGGCG	1024
Db	1004	GCCGCGGCGCTCGACCGCGCGCAAGCGCCCGCGGGGACCGCGCGCGCGCGCGGCG	1063
QY	1025	CACGCGGCTGAGGCGCTGGCGCTCAAGCGGACGCTGACCGAGAGCGGCGAGCG	1084
Db	1064	CACGCGGCTGAGGCGCTGGCGCTCAAGCGGACGCTGACCGAGAGCGGCGAGCG	1123
QY	1085	GCTCCCGGCGCGCGCGCGCGCTGGGCTGGCTGCGCTTCCCGCTGCGGCTGCGGCTGC	1144
Db	1124	GCTCCCGGCGCGCGCGCGCGCTGGGCTGGCTGCGCTTCCCGCTGCGGCTGCGGCTGC	1183
QY	1145	CAGTCCCGCTGCGCTGCTGTGCGGAGCGCGCGCTAGCTCCCGAGGGCGCGCGCGCC	1204
Db	1184	CAGTCCCGCTGCGCTGCTGTGCGGAGCGCGCGCTAGCTCCCGAGGGCGCGCGCGCC	1243
QY	1205	GGACCGAGCGCGCGGACAGAGCAAGGAGGAGGCTGTGCTGGCCACGGCCATCGAGA	1264
Db	1244	GGACCGAGCGCGCGGACAGAGCAAGGAGGAGGCTGTGCTGGCCACGGCCATCGAGA	1303
QY	1265	TCTGGCTCTGA	1275
Db	1304	TCTGGCTCTGA	1314
RESULT 9	AX455339	LOCUS	AX455339
DEFINITION	Sequence 3 from Patent WO0210401.	DNA	linear
ACCESSION	AX455339	1041 bp	PAT 06-JUL-2002
VERSION	AX455339.1		
KEYWORDS	GI:21714512		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

AUTHORS	Meyers, R., Kapeller-Libermann, R. and Silos-Santiago, I.									
TITLE	18431 and 32374, human protein kinase family members and uses therefor									
JOURNAL	Patent: WO 0210401-A 3 07-FEB-2002; Millennium Pharmaceuticals, Inc. (US)									
FEATURES	Location/Qualifiers									
source	1. .1041 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"									
BASE COUNT	147	a	372	c	359	g	163	t		
ORIGIN										
Query Match	81.6%; Score 1041; DB 6; Length 1041;									
Best Local Similarity	100.0%; Pred. No. 3e-123;									
Matches 1041; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;			
QY	235	ATGGCACTGAAGTTTGTGAACAGAGGAAACCAAGCTGAAGAACTTCTTACGGGAGGTG	294							
Db	1	ATGGCACTGAAGTTTGTGAACAGAGGAAACCAAGCTGAAGAACTTCTTACGGGAGGTG	60							
QY	295	AGCATCACCAACAGAGCTTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGTCTTT	354							
Db	61	AGCATCACCAACAGAGCTTCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGTCTTT	120							
QY	355	GAGACAGAGACTGTCTACGTCTTTGCCAGGAGTACGACCTCTGGGACCTGTTTGAC	414							
Db	121	GAGACAGAGACTGTCTACGTCTTTGCCAGGAGTACGACCTCTGGGACCTGTTTGAC	180							
QY	415	ATCATCCTCCCGAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACAGAGCTG	474							
Db	181	ATCATCCTCCCGAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACAGAGCTG	240							
QY	475	GGCTGCGCTGGACTTCATGCAAGGCGGAGTGTGTGCAACGCGACATCAAGCCCGAG	534							
Db	241	GGCTGCGCTGGACTTCATGCAAGGCGGAGTGTGTGCAACGCGACATCAAGCCCGAG	300							
QY	535	AACTGCTGTCTGTGACCGCGAGTGCCTGCGCGCTGAAGCTGGCCGACTTCGGCATCAGC	594							
Db	301	AACTGCTGTCTGTGACCGCGAGTGCCTGCGCGCTGAAGCTGGCCGACTTCGGCATCAGC	360							
QY	595	CGCGCGTGGCTGCGCGCTCAAGCGGCTGAGCGGACCATCCCTTACAGCGCGCTGAG	654							
Db	361	CGCGCGTGGCTGCGCGCTCAAGCGGCTGAGCGGACCATCCCTTACAGCGCGCTGAG	420							
QY	655	GTGTGCCAGGCGGCGCGCGCGCTGCGCGTGTGACAGCGGCGTGTGAGAGCTGTGGGCC	714							
Db	421	GTGTGCCAGGCGGCGCGCGCGCTGCGCGTGTGACAGCGGCGTGTGAGAGCTGTGGGCC	480							
QY	715	TTGCGCGTGTCTCATCTTCTGCGTCTCAACGGCAACTTCCCGTGGGAGGCGGCTCGGGC	774							
Db	481	TTGCGCGTGTCTCATCTTCTGCGTCTCAACGGCAACTTCCCGTGGGAGGCGGCTCGGGC	540							
QY	775	GCCGACGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGGGCGCGCTGCGGGGCTGCT	834							
Db	541	GCCGACGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGGGCGCGCTGCGGGGCTGCT	600							
QY	835	TGCAAGTGGCGCGCTTCAACGAGCGCGCTGTGCGATGTTTCAAGCGCTTACTGGCCCTG	894							
Db	601	TGCAAGTGGCGCGCTTCAACGAGCGCGCTGTGCGATGTTTCAAGCGCTTACTGGCCCTG	660							
QY	895	GAGCCCGAGCGCGGCGGCCAGCCAAAGAGTGTTCGCTTCCCTCAAGCAGAGCTCAGC	954							
Db	661	GAGCCCGAGCGCGGCGGCCAGCCAAAGAGTGTTCGCTTCCCTCAAGCAGAGCTCAGC	720							
QY	955	TCCGAGCTGCGCGCGCGGCTTCGACACGGCGCGCAAGCCCCCGGGGACCGCGCGGCC	1014							
Db	721	TCCGAGCTGCGCGCGCGGCTTCGACACGGCGCGCAAGCCCCCGGGGACCGCGCGGCC	780							
QY	1015	GCCGCGGCGCACTGCGCTCGAGGCGCTGTGGCGCTCAAGCGGAGCGGTGCTGACCGAG	1074							
Db	781	GCCGCGGCGCACTGCGCTCGAGGCGCTGTGGCGCTCAAGCGGAGCGGTGCTGACCGAG	840							

```
Qy 181 AAAGGACCTATGGGAAGTTGACCTGGTGTCTACAAGGGCAGGCAAAAAATGGCA 240
Db 181 AAAGGACCTATGGGAAGTTGACCTGGTGTCTACAAGGGCAGGCAAAAAATGGCA 240
Qy 241 CTGAAGTTTGTGAACAAGACAAACCAAGCTGAAGAACTTCCTACGGGAGGTGAGCATC 300
Db 241 CTGAAGTTTGTGAACAAGACAAACCAAGCTGAAGAACTTCCTACGGGAGGTGAGCATC 300
Qy 301 ACCAAGCCTCTCTCCAGCCCTTCAATCATCAAGGTCCTTTGAGGCTTTTGAGACA 360
Db 301 ACCAAGCCTCTCTCCAGCCCTTCAATCATCAAGGTCCTTTGAGGCTTTTGAGACA 360
Qy 361 GAGGACTGTACGCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTTGAATCATC 420
Db 361 GAGGACTGTACGCTTTTGGCCAGGAGTACGACCTGCTGGGACCTGTTTGAATCATC 420
Qy 421 CCTCCCAAGTGGGCTCCCTGAGGACACGGTGAAGCCTGTGTGACAGCTGGGCTG 480
Db 421 CCTCCCAAGTGGGCTCCCTGAGGACACGGTGAAGCCTGTGTGACAGCTGGGCTG 480
Qy 481 CGCTGGACTTCATGACCGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 540
Db 481 CGCTGGACTTCATGACCGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 540
Qy 541 CTGCTGTTCGACCGGAGTGGCGGCTAAAGCTGGCGGACTTCGGCATACCGGCGGC 600
Db 541 CTGCTGTTCGACCGGAGTGGCGGCTAAAGCTGGCGGACTTCGGCATACCGGCGGC 600
Qy 601 GTGGGCTGCCGCGTCAAGCGGTGAGCGGCACCATCCCTTACAGCGCCCTGAGGTGTC 660
Db 601 GTGGGCTGCCGCGTCAAGCGGTGAGCGGCACCATCCCTTACAGCGCCCTGAGGTGTC 660
Qy 661 CAGCGGGCGCGCGGACGCGGTGCGGTGAGACAGCGGCTGAGAGTGTGGGCTTCGCG 720
Db 661 CAGCGGGCGCGCGGACGCGGTGCGGTGAGACAGCGGCTGAGAGTGTGGGCTTCGCG 720
Qy 721 GTGCTCATCTTCTGGTGTCTCACGGCACTTCCGCTGGGAGGCGGTGCGGCGCCGAC 780
Db 721 GTGCTCATCTTCTGGTGTCTCACGGCACTTCCGCTGGGAGGCGGTGCGGCGCCGAC 780
Qy 781 GCCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGGGCCCTGCGGGGCTCCCTTCGCG 840
Db 781 GCCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGGGCCCTGCGGGGCTCCCTTCGCG 840
Qy 841 TGGCGCGCTTTCACGAGCGCGCTGCGCATGTTCAGGCGCTTACTGGGCGCTTCGAG 900
Db 841 TGGCGCGCTTTCACGAGCGCGCTGCGCATGTTCAGGCGCTTACTGGGCGCTTCGAG 900
Qy 901 GAGCGCGCGCGCGCTTCGACCGCGCGGCAAGCCCGCGGGGACCGCCCGCGCGCGCC 1020
Db 901 GAGCGCGCGCGCGCTTCGACCGCGCGGCAAGCCCGCGGGGACCGCCCGCGCGCGCC 1020
Qy 1021 GGGGCACTGCGGCTGAGGCGCTGGGCGCTCAAGCGGACGCTGCTCAAGCGAGCGGC 1080
Db 1021 GGGGCACTGCGGCTGAGGCGCTGGGCGCTCAAGCGGACGCTGCTCAAGCGAGCGGC 1080
Qy 1081 AGCGCTCCCGCGCGCGCGCTGCGGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1140
Db 1081 AGCGCTCCCGCGCGCGCGCTGCGGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 1140
Qy 1141 GTGCCAGTCCCGCTGCGGCTGCTGCGCGGACCGCGGCTTAGCTTCCCGAGGGGCGCC 1200
Db 1141 GTGCCAGTCCCGCTGCGGCTGCTGCGCGGACCGCGGCTTAGCTTCCCGAGGGGCGCC 1200
Qy 1201 GCGCGGACCGGCGCGCGGACAAAGGCAAGGCGAGGTGCTGCTGCGGACCGGCGCATC 1260
Db 1201 GCGCGGACCGGCGCGCGGACAAAGGCAAGGCGAGGTGCTGCTGCGGACCGGCGCATC 1260
Qy 1261 GAGATCTGGCTC 1272
```

```
Db 1261 GAGATCTGGCTC 1272
RESULT 8
AX455337
LOCUS AX455337 2893 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 1 from Patent WO0210401.
ACCESSION AX455337
VERSION AX455337.1 GI:21714510
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Meyers,R., Kapeller-Libermann,R. and Silos-Santiago,I.
TITLE 18431 and 32374, human protein kinase family members and uses
therefor
JOURNAL Patent: WO 0210401-A 1 07-FEB-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..2893
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
274..1314
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38235.1"
/db_xref="GI:21714511"
/translation="MALKFNKSKTKLKNFLRVSITNSLSSPFFIIKVDVVPETD
CYVFAQBYAPAGDLFDIIPQVGLPFDITKRCVQLGLALDFMHRGLVHRDIKPENV
LLDFRCRVLKADFGMTRVYGVCRVKRVSQTIPTAPEVQAGSADGLAVDTGYDVA
FGLVIFCVLTGNFPWEAASGADAFEFEFVQRGLPGLPQWRFTPEPALRMFQRL
ALPERRGPAKEVFRFLKHELTSCLRPPSHRARKPGDRPPAAGPLRLRAPGLKRT
VLRESGSRPAPPAVGSVPLPVFPVFPVPEPLAPQGPGRITDGRADKSKG
QVLIATAIEICV"
BASE COUNT 500 a 989 c 929 g 466 t 9 others
ORIGIN
```

Query Match 88.8%; Score 1131.8; DB 6; Length 2893;

Best Local Similarity 94.8%; Pred. No. 6.7e-135;

Matches 1205; Conservative 0; Mismatches 57; Indels 9; Gaps 3;

Qy 14 GCCCAGAGCCTGAGCGCCCGCTCCCTGACCTGCTGTGGCGCGGAGCTGCCCTCGGC 73

Db 44 GCCTTGAACCGCCCGCTCCTTGACCTTGTGGCGGAGGAGCTGCCCTCGGC 103

Qy 74 CTGGTGGCG---GTGTGCGCTTCTCACTGAGACATGACGCGCTCACTTCGGACAC 130

Db 104 CTGGTGGCGGTGTGCGCCCTTCTCACTGAAGACATGACGCGCTCACTTCGGACAC 163

Qy 131 TGG--CGCGCAGCAGCTCACCAAGCACTAGCAACTAGT---CCGGAGCTGGGCAAG 184

Db 164 TTGGCCCGCAGCGAGCTCACCCAGCACCTACGAACTAGTCCCGGAGCTGGGCAAG 223

Qy 185 GCACCTATGGAAGTTGACCTGGTGTCTACAAGGGCAGGCAAAAAATGGCACTGA 244

Db 224 GCACCTACGGGAGTTGATCTGGTGTCTACAAGGGCAGGCAAAAAATGGCACTGA 283

Qy 245 AGTTTGAAACAGAGCAAAACCAAGCTGAAGAACTTCCTACGGAGGTGAGCATACCA 304

Db 284 AGTTTGAAACAGAGCAAAACCAAGCTGAAGAACTTCCTACGGAGGTGAGCATACCA 343

Qy 305 ACAGCTCTCTCCAGCCCTTCATCATCAAGGCTTTTGAAGTCTTTGAGACAGG 364

Db 344 ACAGCTCTCTCCAGCCCTTCATCATCAAGGCTTTTGAAGTCTTTGAGACAGG 403

Qy 365 ACTGCTACGCTTTTGCCAGGAGTACGACCTGCTGGGACCTGTTTGACATCATCCTC 424

Db 404 ACTGCTACGCTTTTGCCAGGAGTACGACCTGCTGGGACCTGTTTGACATCATCCTC 463

REFERENCE 1
AUTHORS Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase-like protein
JOURNAL Patent: WO 02083882-A 9 24-OCT-2002;
Bayer Aktiengesellschaft (DE)
FEATURES Location/Qualifiers
source 1.1982
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 323 a 676 c 675 g 308 t
ORIGIN
Query Match 99.9%; Score 1273.4; DB 6; Length 1982;
Best Local Similarity 99.9%; Pred. No. 7.4e-153;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 60
Db 357 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 416
QY 61 ACTGCCCTGGGCTGGTGGCGGTGCGCCCTTCTCACTGAAGACATGACGGCCCTGACT 120
Db 417 ACTGCCCTGGGCTGGTGGCGGTGCGCCCTTCTCACTGAAGACATGACGGCCCTGACT 476
QY 121 CTCGGCACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGGAGCTGGG 180
Db 477 CTCGGCACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGGAGCTGGG 536
QY 181 AAGGCACTATGGAAGGTTGACTGGTGTCTTACAGGGCAAGGCAACAAATGGCA 240
Db 537 AAGGCACTATGGAAGGTTGACTGGTGTCTTACAGGGCAAGGCAACAAATGGCA 596
QY 241 CTGAAGTTGTACACAGACACCAAGCTGAAGACTTCTTACGGAGGTGACATC 300
Db 597 CTGAAGTTGTACACAGACACCAAGCTGAAGACTTCTTACGGAGGTGACATC 656
QY 301 ACCAACAGCTCTCTCCAGCGCCCTTCACTCAAGGTCTTTGAGTCTTTGAGACA 360
Db 657 ACCAACAGCTCTCTCCAGCGCCCTTCACTCAAGGTCTTTGAGTCTTTGAGACA 716
QY 361 GAGACTGCTACGCTTTTGGCAGAGTACGACCTGCTGGGGACCTTTTGACATCATC 420
Db 717 GAGACTGCTACGCTTTTGGCAGAGTACGACCTGCTGGGGACCTTTTGACATCATC 776
QY 421 CCTCCAGAGTGGGCTCCCTGAGACACGGTGAAGCGCTGTGTGACAGCTGGGCGCTG 480
Db 777 CCTCCAGAGTGGGCTCCCTGAGACACGGTGAAGCGCTGTGTGACAGCTGGGCGCTG 836
QY 481 GCGCTGGAATTCATGACGGGCGGAGCTGTGTGACACCGGACATCAAGCCGAGAACCTG 540
Db 837 GCGCTGGAATTCATGACGGGCGGAGCTGTGTGACACCGGACATCAAGCCGAGAACCTG 896
QY 541 CTGCTGTTGACCGGAGTGGCGCGCTTAAGCTGGCGGCTTCCGATGACGCGCGCG 600
Db 897 CTGCTGTTGACCGGAGTGGCGCGCTTAAGCTGGCGGCTTCCGATGACGCGCGCG 956
QY 601 GTGGCTGCGCGCTCAAGCGCGTGAAGCGGACCATCCCTTACACGGCGCTGAGGTGTC 660
Db 957 GTGGCTGCGCGCTCAAGCGCGTGAAGCGGACCATCCCTTACACGGCGCTGAGGTGTC 1016
QY 661 CAGGCGGCGCGCGGACCGGCTGGCGGTGGACACGCGGCTGGAAGTGGGCGCTTCGCG 720
Db 1017 CAGGCGGCGCGCGGACCGGCTGGCGGTGGACACGCGGCTGGAAGTGGGCGCTTCGCG 1076
QY 721 GTGCTCATTTCTGGTGTCTACCGGCACTTCCCGTGGGAGCGCGCTTCGGGCGCGAC 780
Db 1077 GTGCTCATTTCTGGTGTCTACCGGCACTTCCCGTGGGAGCGCGCTTCGGGCGCGAC 1136
QY 781 GCCTTCTTCGAGAGTTCGTGGCTGGCAGCGGGGCGCGCTTCGGGGGTGCTTCGAG 840
Db 1137 GCCTTCTTCGAGAGTTCGTGGCTGGCAGCGGGGCGCGCTTCGGGGGTGCTTCGAG 1196

QY 841 TGGCGCGGCTTCAAGAGCCCGCGCTGGCGATGTTTCAAGCGCTTACTGGCCCTTGAGCCCC 900
Db 1197 TGGCGCGGCTTCAAGAGCCCGCGCTGGCGATGTTTCAAGCGCTTACTGGCCCTTGAGCCCC 1256
QY 901 GAGCGCGGCGGCGCCAGCAAGAGAGTGTCCCGCTTCTCAAGCAGAGCTCACGTCCGAG 960
Db 1257 GAGCGCGGCGGCGCCAGCAAGAGAGTGTTCCTTCTCAAGCAGAGCTCACGTCCGAG 1316
QY 961 CTGCGCGCGCGGCTTCAAGAGCCCGCGCTCAAGCGGCGGAGACCGCGCGCGCGCG 1020
Db 1317 CTGCGCGCGCGGCTTCAAGAGCCCGCGCTCAAGCGGCGGAGACCGCGCGCGCGCG 1376
QY 1021 GGGCCACTGCGGCTTCAAGAGCCCGCGCTCAAGCGGCGGAGACCGGTGCTGACCGAGCGCG 1080
Db 1377 GGGCCACTGCGGCTTCAAGAGCCCGCGCTCAAGCGGCGGAGACCGGTGCTGACCGAGCGCG 1436
QY 1081 AGCGGCTCCCG 1140
Db 1437 GGGCGCTCCCG 1496
QY 1141 GTCCAGTGGCGGCTGCGGCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Db 1497 GTCCAGTGGCGGCTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1556
QY 1201 GCGCGAGCGAGCG 1260
Db 1557 GCGCGAGCGAGCG 1616
QY 1261 GAGATCTGCGCTCTGA 1275
Db 1617 GAGATCTGCGCTCTGA 1631
RESULT 7
AX589472 1272 bp DNA linear PAT 24-JAN-2003
LOCUS Sequence 1 from Patent WO20083882.
DEFINITION AX589472
ACCESSION AX589472
VERSION AX589472.1 GI:27900910
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase-like protein
JOURNAL Patent: WO 02083882-A 1 24-OCT-2002;
Bayer Aktiengesellschaft (DE)
FEATURES Location/Qualifiers
source 1.1272
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 191 a 450 c 430 g 201 t
ORIGIN
Query Match 99.0%; Score 1262.4; DB 6; Length 1272;
Best Local Similarity 99.5%; Pred. No. 2.1e-151;
Matches 1266; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 60
Db 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 60
QY 61 ACTGCCCTGGGCTGGTGGCGGTGCGCCCTTCTCACTGAAGACATGACGGCCCTGACT 120
Db 61 ACTGCCCTGGGCTGGTGGCGGTGCGCCCTTCTCACTGAAGACATGACGGCCCTGACT 120
QY 121 CTCGGCACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGGAGCTGGG 180
Db 121 CTCGGCACACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGGAGCTGGG 180

QY	661	CAGGCGGCGCGCGGCTGGCGGTGGACACGCGCGTGGAGCTGTGGGCTTCGGC	720
Db	1434	CAGGCGGCGCGCGGCTGGCGGTGGACACGCGCGTGGAGCTGTGGGCTTCGGC	1493
QY	721	GTGCTCATCTTCTGGTGTCTACCGGCAACTTCCCGTGGAGGCGGCTCGGGCGCGAC	780
Db	1494	GTGCTCATCTTCTGGTGTCTACCGGCAACTTCCCGTGGAGGCGGCTCGGGCGCGAC	1553
QY	781	GCTTCTTCGAGAGTCTTGTGGCTGGCAGCGGCGCGCTTGCCTGGGCTGCTTCGCGAG	840
Db	1554	GCTTCTTCGAGAGTCTTGTGGCTGGCAGCGGCGCGCTTGCCTGGGCTGCTTCGCGAG	1613
QY	841	TGGCGCGCTTCAACGAGCCCGCTGCGCATGTTTCCAGGCGTTACTGGCGCTGGAGGCC	900
Db	1614	TGGCGCGCTTCAACGAGCCCGCTGCGCATGTTTCCAGGCGTTACTGGCGCTGGAGGCC	1673
QY	901	GAGCGCGCGCGCGCTCGCACCGCGCGGCAAGAGGTTTCCGCTTCTTCAAGCAGAGCTCAGTCCGAG	960
Db	1674	GAGCGCGCGCGCGCTCGCACCGCGCGGCAAGAGGTTTCCGCTTCTTCAAGCAGAGCTCAGTCCGAG	1733
QY	961	CTGGCGCGCGCGCTCGCACCGCGCGGCAAGAGGTTTCCGCTTCTTCAAGCAGAGCTCAGTCCGAG	1020
Db	1734	CTGGCGCGCGCGCTCGCACCGCGCGGCAAGAGGTTTCCGCTTCTTCAAGCAGAGCTCAGTCCGAG	1793
QY	1021	GGGCGACTGCGCTCGAGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	1080
Db	1794	GGGCGACTGCGCTCGAGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	1853
QY	1081	AGCGGCTCCCGCGCGCGCGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	1140
Db	1854	AGCGGCTCCCGCGCGCGCGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	1913
QY	1141	GTGCGAGTCCCGCGCGCGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	1200
Db	1914	GTGCGAGTCCCGCGCGCGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	1973
QY	1201	GGCGGACCGAGCGCGCGCGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	1260
Db	1974	GGCGGACCGAGCGCGCGCGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGGC	2033
QY	1261	GAGATCTGGTCTGA	1275
Db	2034	GAGATCTGGTCTGA	2048
RESULT 4			
LOCUS	AX166553	1275 bp	DNA linear PAT 22-JUN-2001
DEFINITION	Sequence 44 from Patent WO0138503.		
ACCESSION	AX166553		
VERSION	AX166553.1	GI:14546898	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. Plowman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S.S., Martinez, R., Flanagan, P. and Clary, D.S.		
TITLE	Novel human protein kinases and protein kinase-like enzymes		
JOURNAL	Patent: WO 0138503-A 44 31-MAY-2001;		
FEATURES	Sugen, Inc. (US)		
source	Location/Qualifiers		
	1. .1275		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
BASE COUNT	191 a 450 c 434 g 200 t		
ORIGIN			
Query Match	99.9%;	Score 1273.4;	DB 6; Length 1275;
Best Local Similarity	99.9%;	Pred. No. 8.3e-153;	

Matches 1274;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
1 QY	1 ATGAGCGTGGGCTGCCAGAGCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG	60		
1 Db	1 ATGAGCGTGGGCTGCCAGAGCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG	60		
61 QY	61 ACTGCCCTGGGCTGGTGGGCTGTGCCCTTCTCACTGAAGACATGACGAGCCCTGACT	120		
61 Db	61 ACTGCCCTGGGCTGGTGGGCTGTGCCCTTCTCACTGAAGACATGACGAGCCCTGACT	120		
121 QY	121 CTCCGACACTGGCGCGCAGCGAGCTCAACAAGCACTACGAACCTAGTCCGGGAGCTGGGC	180		
121 Db	121 CTCCGACACTGGCGCGCAGCGAGCTCAACAAGCACTACGAACCTAGTCCGGGAGCTGGGC	180		
181 QY	181 AAAGGCACCTATGGGAAGTTGACCTGTGTGTCTTCAAGGGCAAGGCACAAAAATGGCA	240		
181 Db	181 AAAGGCACCTATGGGAAGTTGACCTGTGTGTCTTCAAGGGCAAGGCACAAAAATGGCA	240		
241 QY	241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGGAGGTGAGCATC	300		
241 Db	241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGGAGGTGAGCATC	300		
301 QY	301 ACCAAACAGCTCTCTCTCCAGCGCCCTTCAATCAATCAAGGTCTTTGACGTGGTCTTTGAGACA	360		
301 Db	301 ACCAAACAGCTCTCTCTCCAGCGCCCTTCAATCAATCAAGGTCTTTGACGTGGTCTTTGAGACA	360		
361 QY	361 GAGACTGTCTACGCTTTTGGCCAGGAGTACGCACCTGTCTGGGACCTCTTTTGACATCATC	420		
361 Db	361 GAGACTGTCTACGCTTTTGGCCAGGAGTACGCACCTGTCTGGGACCTCTTTTGACATCATC	420		
421 QY	421 CCTCCCGAGGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGCGAGCAGCTGGGCTG	480		
421 Db	421 CCTCCCGAGGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGCGAGCAGCTGGGCTG	480		
481 QY	481 GCGCTGACTTTCATGCAACGCGCGCAGCTGTGTGTGCAACCGGACATCAAGCCCGAGAACCTG	540		
481 Db	481 GCGCTGACTTTCATGCAACGCGCGCAGCTGTGTGTGCAACCGGACATCAAGCCCGAGAACCTG	540		
541 QY	541 CTGCTGTTTCGACCGGAGTGCCCGCGGTAAAGCTGGCCGACTTCGGCATACGCGCGCGC	600		
541 Db	541 CTGCTGTTTCGACCGGAGTGCCCGCGGTAAAGCTGGCCGACTTCGGCATACGCGCGCGC	600		
601 QY	601 GTGGGCTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGCGCTTACGAGTGTGC	660		
601 Db	601 GTGGGCTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGCGCTTACGAGTGTGC	660		
661 QY	661 CAGCGGCGCGCGCGGCTGGCGGTGGACACGGGCGTGGAGCTGTGGGCGCTTCGGC	720		
661 Db	661 CAGCGGCGCGCGCGGCTGGCGGTGGACACGGGCGTGGAGCTGTGGGCGCTTCGGC	720		
721 QY	721 GTGCTCATCTTCTGGTGTCTACCGGCAACTTCCCGTGGGAGCGGCTCGGGGCGCGAC	780		
721 Db	721 GTGCTCATCTTCTGGTGTCTACCGGCAACTTCCCGTGGGAGCGGCTCGGGGCGCGAC	780		
781 QY	781 GCCTTCTTCGAGAGTTTGTGCGCTGGAGCGGCGCGCTTCCAGCGGCTTACTGGGAGTGCCTTCGAG	840		
781 Db	781 GCCTTCTTCGAGAGTTTGTGCGCTGGAGCGGCGCGCTTCCAGCGGCTTACTGGGAGTGCCTTCGAG	840		
841 QY	841 TGGCGCGCTTCAACGAGCCCGCGCTGCGCATGTTTCCAGCGCTTACTGGCCCTGGAGGCC	900		
841 Db	841 TGGCGCGCTTCAACGAGCCCGCGCTGCGCATGTTTCCAGCGCTTACTGGCCCTGGAGGCC	900		
901 QY	901 GAGCGCGCGCGCGCAGCAAGGAGGTGTTCGCTTCTCAAGCAGAGCTCAGCTTCGAG	960		
901 Db	901 GAGCGCGCGCGCGCAGCAAGGAGGTGTTCGCTTCTCAAGCAGAGCTCAGCTTCGAG	960		
961 QY	961 CTGGCGCGCGCGCTTCGACACCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1020		
961 Db	961 CTGGCGCGCGCGCTTCGACACCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1020		
1021 QY	1021 GGGCGACTGCGGCTCGAGGCGGCTTGGGCGGCTCAAGCGGACGGTGTCTACCGAGAGCGGC	1080		
1021 Db	1021 GGGCGACTGCGGCTCGAGGCGGCTTGGGCGGCTCAAGCGGACGGTGTCTACCGAGAGCGGC	1080		

241 CTGAGTTTGTCAACAGCAGCAGACCAAGCTGAGAGCTTCCTACCGGAGGTGAGCATC 300
|||
357 CTGAAGTTTGTGAACAGAGCAAAACCAAGCTGAAGACTTCCTACCGGAGGTGAGCATC 416
|||
301 ACCAACAGCCTCTCTCCAGCCCTTCATCAAGGTCTTTGACGTGCTTTTGAGACA 360
|||
417 ACCAACAGCCTCTCTCCAGCCCTTCATCAAGGTCTTTGACGTGCTTTTGAGACA 476
|||
361 GAGGACTGCTAGCTTTTCCAGAGTACGACCTCTGTTGGGACCTGTTGACATCATC 420
|||
477 GAGGACTGCTAGCTTTTCCAGAGTACGACCTCTGTTGGGACCTGTTGACATCATC 536
|||
421 CTTCCCGAGGTGGGACTCCCTGAGACACAGGTGAAGCTGTGTGACAGCTGGGCTG 480
|||
537 CTTCCCGAGGTGGGACTCCCTGAGACACAGGTGAAGCTGTGTGACAGCTGGGCTG 596
|||
481 GCGCTGGACTTCATGACAGGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 540
|||
597 GCGCTGGACTTCATGACAGGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 656
|||
541 CTGCTGTTCGACCGGAGTCCCGCGGTAAAGCTGGCGGACTTCGCGATGACGCGCGC 600
|||
657 CTGCTGTTCGACCGGAGTCCCGCGGTAAAGCTGGCGGACTTCGCGATGACGCGCGC 716
|||
601 GTGGCTGCGCGGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTC 660
|||
717 GTGGCTGCGCGGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTC 776
|||
661 CAGCGCGCGCGCGGTGCGGTGAGCGGTGAGCGGCTGAGCGGTGAGCGGTGAGCGGTG 720
|||
777 CAGCGCGCGCGCGGTGCGGTGAGCGGTGAGCGGTGAGCGGTGAGCGGTGAGCGGTG 836
|||
721 GTGCTCATCTTCTGCTGCTCAACGCGACTTCCGCTGGGAGCGGCTGGGCGCGCGAC 780
|||
837 GTGCTCATCTTCTGCTGCTCAACGCGACTTCCGCTGGGAGCGGCTGGGCGCGCGAC 896
|||
781 GCTTCTTCGAGAGTTCTGCTGCTGGAGCGGCGCGCTCCCGGGGTGCTTCGCGAG 840
|||
897 GCTTCTTCGAGAGTTCTGCTGCTGGAGCGGCGCGCTCCCGGGGTGCTTCGCGAG 956
|||
841 TGGCGCGCTTCCAGAGCGCGCTGCGCATGTTCCAGCGCTTACTTGGCCCTTGAGGCC 900
|||
957 TGGCGCGCTTCCAGAGCGCGCTGCGCATGTTCCAGCGCTTACTTGGCCCTTGAGGCC 1016
|||
901 GAGCGCGCGCGCGCGAGAGGTGTTCCGCTTCTCAAGCAAGCTACGTCAGTCGCGAG 960
|||
1017 GAGCGCGCGCGCGCGAGAGGTGTTCCGCTTCTCAAGCAAGCTACGTCAGTCGCGAG 1076
|||
961 CTGCGCGCGCGCGCTCGACCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
|||
1077 CTGCGCGCGCGCGCTCGACCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1136
|||
1021 GGGCCACTCTGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGACGCTGTGACCGAGAGCGCG 1080
|||
1137 GGGCCACTCTGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGACGCTGTGACCGAGAGCGCG 1196
|||
1081 ACGGCTCTCCG 1140
|||
1197 ACGGCTCTCCG 1256
|||
1141 GTGCCAGTCTCCGCTGCTGTGCGCGAGCGCGCGCTAGCTCCCGAGGGCGCGCGCG 1200
|||
1257 GTGCCAGTCTCCGCTGCTGTGCGCGAGCGCGCGCTAGCTCCCGAGGGCGCGCGCG 1316
|||
1201 GGGCGGACCGAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
|||
1317 GGGCGGACCGAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1376
|||
1261 GAGATCTGCTGTGA 1275
|||
1377 GAGATCTGCTGTGA 1391
|||

RESULT 3
AX364846
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2598
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 446 a 952 c 862 g 338 t
ORIGIN
Query Match 100.0%; Score 1275; DB 6; Length 2598;
Best Local Similarity 100.0%; Pred.No. 4.4e-153;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGCTGCCAGAGCTGAGCGCGCCCTCCCTGACCTGCTGTGGGCGGGG 60
DB 774 ATGAGCGTGGCTGCCAGAGCTGAGCGCGCCCTCCCTGACCTGCTGTGGGCGGGG 833
QY 61 ACTGCCCTGGGCTGGTGGCGGTGTGCGCTTCTCACTGAAGACATCAGGCCCTGACT 120
DB 834 ACTGCCCTGGGCTGGTGGCGGTGTGCGCTTCTCACTGAAGACATCAGGCCCTGACT 893
QY 121 CTCGCGACACTGGCGCGCAGGAGCTGACCGAGCTACCGAGCTAGCTAGTCTGGGAGTGGC 180
DB 894 CTCGCGACACTGGCGCGCAGGAGCTGACCGAGCTAGCTAGTCTGGGAGTGGC 953
QY 181 AAAGCACCTATGGAAGGTTGACCTGTGTGTCTCAAGGGCAGAGCACAAAATGCA 240
DB 954 AAAGCACCTATGGAAGGTTGACCTGTGTGTCTCAAGGGCAGAGCACAAAATGCA 1013
QY 241 CTGAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACCGGAGGTGAGCATC 300
DB 1014 CTGAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACCGGAGGTGAGCATC 1073
QY 301 ACCAACAGCCTCTCTCCAGCCCTTCATCAAGGTCTTTGACGTGCTTTTGAGACA 360
DB 1074 ACCAACAGCCTCTCTCCAGCCCTTCATCAAGGTCTTTGACGTGCTTTTGAGACA 1133
QY 361 GAGGACTGCTAGCTTTTCCAGAGTACGACCTGCTGGGACCTGTTGACATCATC 420
DB 1134 GAGGACTGCTAGCTTTTCCAGAGTACGACCTGCTGGGACCTGTTGACATCATC 1193
QY 421 CTTCCCGAGGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACAGCTGGGCTG 480
DB 1194 CTTCCCGAGGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACAGCTGGGCTG 1253
QY 481 GGGCTGGACTTCATGACAGGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 540
DB 1254 GGGCTGGACTTCATGACAGGGCGGAGCTGTGTGACCGGACATCAAGCCCGAGAACGTG 1313
QY 541 CTGCTGTTGACCGGAGTGGCGCGGTAAAGCTGGCGGACTTCGGCATGACGCGCGC 600
DB 1314 CTGCTGTTGACCGGAGTGGCGCGGTAAAGCTGGCGGACTTCGGCATGACGCGCGC 1373
QY 601 GTGGGCTGCCGCTCAAGCGGTGAGCGGCGACCATCCCTTACACGCGCGCTGAGGTGTC 660
DB 1374 GTGGGCTGCCGCTCAAGCGGTGAGCGGCGACCATCCCTTACACGCGCGCTGAGGTGTC 1433
|||

[illegible]

595	TTCTGTACAGTTCTGTGTGGGCTCTGCTGGGGCGGCACATGAGCCATCGTGCAGGCC	654
340	TTTGACGTGCTTTTGAGACAGAGGACTGCTAGTCTTTTCCAGGAGTAGTACGCACCTGCT	399
655	TACGGCATTTGGCATCGAGTCGGCACACTCTACAGCTTCTGTACGGAGCCGCTCTGCAC	714
400	GGGGACGTGTTGACATCATCCCTCCAGGTGGGCTCCCTGAGGACACAGGTGAGACGC	459
715	GGGGACCTCATGGCTTTCATCCAGCCAAAGGTGGGCTCCCGCAGCCGCGGGTGCACCGC	774
460	TGTGTGACAGAGTGGGCTCGGGCTTGAGCTTCATGCACGGGGCGGAGCTGSGTGCACCGC	519
775	TGCGCCGCCAGCTGGCTCCGCCCTTGGAGTATACATCCACGCCCGGSCCTGGTGTACCGG	834
520	GACATCAAGCCCGAGACGTCTCTCTTTGACCGCGAGTGGCCCGCTTAAAGCTTGCC	579
835	GACCTGAAGCCCGAAGACGTCTGGTGTGGACCCCGGCTTCGCGGGCGCTTCAGAGTGAAC	894
580	GACTTCGGCATGACGCGCCGCTGGGCTGGCGGTCAAGCGCGGTGAGCGGCACATCCCT	639
895	GACTTCGGCCACACAGAGGCTTCGGGGACGCTGCTGCGCTTGCGCGGCGCCATCCCC	954
640	TACACGGGCGCTGAGTGTGC---CAGGGGGCCCGCCGACGGGCTGGCGGTGGACAG	696
955	TACACGGCCCCGAGCTCTGGCGCCCGCCCGCTCCCGAGGCGCTGCGCATTCAGCCC	1014
697	GGCGTGGAAGTGTGGGCTTCGGGCGTGTCTATCTTTCGCGTGTCTACCGGCAACTCCCC	756
1015	GCCCTGAGCGCTGGCGCTGGGGTCTGCTCTTCGCTCTCTACAGGCTACTTCCCC	1074
757	TGGGA---GGCGGCGTGGGGCCGAGCCCTTCTCGAGGAGTTCGTGCGCTGGCAGCGG	813
1075	TGGGACCGGGCCCTGGCGGAGGCGGACCCCTTCTACGAGAGCTTCCTCATCTGGCAGGCG	1134
814	GGCGCGCTTCGCGGGCTGCCTTTCGAGTGGCGCGCTTCACCGAGCCCGGCTGCGCATG	873
1135	TCGGGCGCAGCCCGGAGACCGCCCTCAGCCCTGGTTCGGCTTGGCCCCCGGCGCAGCG	1194
874	TTCAGGCGCTTACTG	888
1195	CTTCTGGGGGCTG	1209

Search completed: December 7, 2003, 15:17:47
Job time : 303 secs

CC e.g. inflammation, hyperalgesia. The invention is also useful in
CC chromosome mapping, tissue typing and in forensic identification of a
CC biological sample. 5884 molecules are useful as markers of disorders or
CC disease states, as markers of drug activity, or as markers of the
CC pharmacogenomic profile of the subject. The present nucleic acid sequence
CC encodes the human novel protein kinase 5884 protein of the invention.

XX
SQ Sequence 1247 BP; 175 A; 454 C; 421 G; 197 T; 0 other;

Query Match 18.6%; Score 236.6; DB 24; Length 1247;
Best Local Similarity 57.9%; Pred. No. 5.5e-31;
Matches 460; Conservative 0; Mismatches 329; Indels 6; Gaps 2;

QY 100 GAAGACATGAGGCGCTGACTCTCCGACACTGCGCCGACGAGCTGACCAAGCACTAC 159
DB |||||
QY 170 GAGGACATGATGACCTGAGTGTCTCAGACCTGTTCGAGCCGAGGTGGAGGCTCTAC 229
DB |||||
QY 160 GAACCTAGTCCGGAGCTGGGCAAGGACCTATGGGAAGTTGACCTGGTGGTCTTACAAG 219
DB |||||
QY 230 GAGGAAGTGGTCCCTGGGCGAGGCTCGCTATGGCGCGTCTCTTGTGTCAACCATCGT 289
DB |||||
QY 220 GGCACAGGCAAAATGGCACTGAAGTTTGTGACAGAGCAAAACCAAGCTGAAGAAC 279
DB |||||
QY 290 CAGAAAGGCACACCCCTGGCACTGAAGAGCTCCCGAAACCCCGCACTGCTCCCTGGC 349
DB |||||
QY 280 TTCCTACGGGAGTGAGCATCACCAAGAGCTCTCTCCAGCCCTTTCATCATCAAGGTC 339
DB |||||
QY 350 TTCCTGACGAGTCTGTGTGGGCTCTCGCTGGCGCGCACTCAGCCATCGTACGGCC 409
DB |||||
QY 340 TTTGACCTGCTTTTGACAGAGAGCTGTAAGTCTTTTGGCCAGGAGTACGACCTGCT 399
DB |||||
QY 410 TACGGCATTTGATCGAGTGGCACTCTTACAGCTTCTTGAAGGAGCCGCTCTCGAC 469
DB |||||
QY 400 GGGGACCTGTTGACATCATCTCCCGAGTGGGCTCCCTGAGGACAGCTGAGCGC 459
DB |||||
QY 470 GGGGACCTCATTTGAGCTTCTCAGCCCAAGTGGGCTTCCCGAGCCCGCGGTGACCGC 529
DB |||||
QY 460 TGTGTGACGAGCTGGGCTGGCGCTGAGCTTCTATGACGCGGCGGAGCTGGTCAACGC 519
DB |||||
QY 530 TGGCGCGGAGCTGGCTTCCGCTTGGAGTACATCCACGCGCGGCTGGTGTACCGG 589
DB |||||
QY 520 GACATCAAGCCCGAAGCTGCTGTGTTGACCGGAGTGGCGCGGCTAAAGTGGCC 579
DB |||||
QY 590 GACCTGAAGCCGGAAGAGCTCTGTGTGTCGACCCCGGCTGCGCGGCTTCAAGCTGACC 649
DB |||||
QY 580 GACTTCGGCATGACGCGCGCTGGGCTGCGCTCAAGCGGCTGAGCGGCAACCATCGCT 639
DB |||||
QY 650 GACTTCGGCCACACGAGGCTTCGCGGAGCTGTGCGCTGGCGCGGCGGCGGCTTCCCG 709
DB |||||
QY 640 TACAGCGGCTGAGGTGTC---CAGGCGGCGCGCGCGGCTGGCGGTGGACACG 696
DB |||||
QY 710 TACAGGCGCGGAGCTGTGCGGCGCGCGCGCTCCCGAGGCGCTGCGCATTCAGGCC 769
DB |||||
QY 697 GCGGTGGAGCTGGGCTTGGCGGTGCTATCTTCTGGGTGCTACCGGCACTTCGG 756
DB |||||
QY 770 GCGCTGGAGCGCTGGGCGCTGGGCTGCTCTTCTGCTCTCTCAAGGCTTCTCCCG 829
DB |||||
QY 757 TGGGA---GGCGGCTGGGCGCGGAGCTTCTTGGAGAGTGTGCGGTGGAGCGG 813
DB |||||
QY 830 TGGGACCGGCGGCTGGCGGAGGCGGAGCTTCTTACGAGACTTCTCTATCTGGCAGGCG 889
DB |||||
QY 814 GGCCTGCTGCGGCGGCTGCTTTCGAGTGGCGGCTTCAAGGAGCGCGGCTGGCGATG 873
DB |||||
QY 890 TCGGGCCAGCCCGGAGCGGCTTCTGAGCTTCTGAGGAGTGTGCGGTGGAGCGG 949
DB |||||
QY 874 TTCAGCGGTACTG 888
DB |||||
QY 950 CTTCTGGCGGGGCTG 964
DB |||||

RESULT 15
AAS06706
ID AAS06706 standard; cdna; 1716 BP.

XX
AC AAS06706;
XX 12-SEP-2001 (first entry)
XX
DE Polynucleotide sequence encoding human protein kinase #6.
XX
XX Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
XX reproductve disorder; gene therapy; ss.
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US32085.
XX
XX 24-NOV-1999; 99US-0167482.
XX
XX (SUGB-) SUGEN INC.
XX
XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX
XX WPI; 2001-343950/36.
XX
XX P-PSDB; AAU03506.
XX
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
XX Example 1; Figure 1; 433pp; English.
XX
XX AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (TK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
XX modulators of protein kinase expression and activity.

XX
SQ Sequence 1716 BP; 285 A; 597 C; 564 G; 270 T; 0 other;

Query Match 18.6%; Score 236.6; DB 22; Length 1716;
Best Local Similarity 57.9%; Pred. No. 5.5e-31;
Matches 460; Conservative 0; Mismatches 329; Indels 6; Gaps 2;

QY 100 GAAGACATGAGGCGCTGACTCTCCGACACTGCGCCGACGAGCTGACCAAGCACTAC 159
DB |||||
QY 415 GAGGACATGATGACCTGAGTGTCTCAGACCTGTTCGAGCCGAGGTGGAGGCTCTAC 474
DB |||||
QY 160 GAACCTAGTCCGGAGCTGGGCAAGGACCTATGGGAAGTTGACCTGGTGGTCTTACAAG 219
DB |||||
QY 475 GAGGAAGTGGTCCCTGGGCGAGGCTCGCTATGGCGCGTCTCTTGTGTCAACCATCGT 534
DB |||||
QY 220 GGCACAGGCAAAATGGCACTGAAGTTTGTGACAGAGCAAAACCAAGCTGAAGAAC 279
DB |||||
QY 535 CAGAAAGGCACACCCCTGGCACTGAAGAGCTCCCGAAACCCCGCACTGCTCCGTGGC 594
DB |||||
QY 280 TTCCTACGGGAGTGAGCATCACCAAGCTCTCTCCAGCGCCCTTTCATCATCAAGGTC 339
DB |||||

CC gene whose imprinting is lost in cancer, or if the test substance
CC inhibits imprinting of a gene whose imprinting is gained in cancer. The
CC methylated CpG islands are useful for providing an assessment of the risk
CC of developing cancer, or for providing diagnostic information relative to
CC cancer which involves determining the methylation status of the CpG
CC island in a patient's DNA. The EG cells allow the accession of imprinted
CC genes which are useful for detecting birth defects, diabetes and cancers
CC associated with aberrant imprinting. The EG cell lines represent the
CC first in vitro model system in which genomic imprinting can be followed
CC dynamically and the two alleles can be distinguished. AAS21030-AAS21083
CC represent isolated CpG islands which are not present in public databases.
XX
XX
SQ Sequence 788 BP; 111 A; 291 C; 300 G; 86 T; 0 other;

Query Match 60.2%; Score 767.6; DB 24; Length 788;
Best Local Similarity 99.4%; Pred. No. 8e-120;
Matches 781; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 420 CCCTCCCCAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGCAGCAGCTGGGCGCT 479
DB 785 CGGCCCGAGGTGGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGCAGCAGCTGGGCGCT 726
QY 480 GCGCTGTGACTTCATGCA CGGCGCGCAGCTGTGTGCACCGCACATCAAGCCCGAGAAAGT 539
DB 725 GCGCTGTGACTTCATGCA CGGCGCGCAGCTGTGTGCACCGCACATCAAGCCCGAGAAAGT 666
QY 540 GCTGCTGTTCACCGCGAGTGC CGCGCGGTAAAGCTGCGCGACTTCGGCATGACCGCGCG 599
DB 665 GCTGCTGTTCACCGCGAGTGC CGCGCGGTAAAGCTGCGCGACTTCGGCATGACCGCGCG 606
QY 600 CGTGGGCTCGCGCTCAAGCGCGTGAAGCGCACCATTCCTTACACGGCGCTGTGAGTGTG 659
DB 605 CGTGGGCTCGCGCTCAAGCGCGTGAAGCGCACCATTCCTTACACGGCGCTGTGAGTGTG 546
QY 660 CCAAGCGCGCGCGCGCGCGCGTGGCGGTGACACCGCGCGTGTGCGGCTTCGG 719
DB 545 CCAAGCGCGCGCGCGCGCGCGTGGCGGTGACACCGCGCGTGTGCGGCTTCGG 486
QY 720 CGTGTGATCTTCGCTGCTCA CGCGCGCACTTCCTCCGTGGAGCGCGTGTGGGCGCGCA 779
DB 485 CGTGTGATCTTCGCTGCTCA CGCGCGCACTTCCTCCGTGGAGCGCGTGTGGGCGCGCA 427
QY 780 CGCTCTTCCTGAGAGTTCGTGCGCTGCGAGCGGCGCGCTTCGCGGCGCTTCGCA 839
DB 426 CGCTCTTCCTGAGAGTTCGTGCGCTGCGAGCGGCGCGCTTCGCGGCGCTTCGCA 367
QY 840 GTGGCGCGCTTCACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTTGAGCC 899
DB 366 GTGGCGCGCTTCACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTTGAGCC 307
QY 900 CGAGCGCGCGCGCGCGCGAGGAGTTCCTGCTTCTCAAGCAGAGCTCAGTCCGA 959
DB 306 CGAGCGCGCGCGCGCGCGAGGAGTTCCTGCTTCTCAAGCAGAGCTCAGTCCGA 247
QY 960 GCTGGCGCGCGCGCGCTCGCACCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCG 1019
DB 246 GCTGGCGCGCGCGCGCTCGCACCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCG 187
QY 1020 CGGGCCACTGCGCTTCGAGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGCGG 1079
DB 186 CGGGCCACTGCGCTTCGAGCGCGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGCGG 127
QY 1080 CAGCGGCTCCCGCGCGCGCGCGCGCGTGGGTGCGTTCGCTGCGCTGCGCGTCCCGTCC 1139
DB 126 CAGCGGCTCCCGCGCGCGCGCGCGCGTGGGTGCGTTCGCTGCGCTGCGCGTCCCGTCC 67
QY 1140 GGTGCGAGTGGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1199
DB 66 GGTGCGAGTGGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 7
QY 1200 CGCGCG 1205
DB 6 CGCGCG 1

RESULT 14
ABK89200
ID ABK89200 standard; DNA; 1247 BP.
XX
AC ABK89200;
XX
DT 21-OCT-2002 (first entry)
XX
DE DNA encoding human novel protein kinase 58848 protein.
XX
KW Human; protein kinase 58848; protein kinase associated disorder;
KW 58848-associated disorder; cellular disorder; proliferative disorder;
KW differentiate disorder; cancer; leukaemia; hormonal disorder; diabetes;
KW immune disorder; autoimmune disease; blood vessel disorder; hypertension;
KW atherosclerosis; platelet disorder; cardiovascular disorder; ischaemia;
KW cardiac hypertrophy; heart failure; neurological disorder; rickets;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; AIDS;
KW acquired immunodeficiency syndrome; bone metabolism disorder; pain;
KW osteoporosis; cirrhosis; haematopoietic neoplastic disorder; Hepatitis B;
KW Hodgkin's disease; acute leukaemia; liver disorders; Gaucher's disease;
KW viral disease; metabolic disorder; inflammation; hyperalgesia;
KW chromosome mapping; tissue typing; forensic identification;
KW pharmacogenomic profile; gene; ds.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FT 1..43
FT 5'UTR /tag= a
FT CDS 44..1090
FT /tag= b
FT /product= "Human novel protein kinase 58848 protein"
FT /note= "Specifically claimed in claim 1"
FT 3'UTR 1091..1247
FT /tag= c
XX
PN WO200255713-A2.
XX
PD 18-JUL-2002.
XX
PF 26-NOV-2001; 2001WO-US44346.
XX
PR 08-DEC-2000; 2000US-254401P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Kapeller-Libermann R, Acton S;
XX
DR WPI; 2002-590676/63.
DR P-PSDB; ABG30935.
XX
PT Novel human protein kinase polypeptide, designated 58848, useful for
PT treating diseases including cellular, proliferative, bone metabolism,
PT cardiovascular, neurological, and haematopoietic neoplastic disorders -
XX Claim 1; Page 97-99; 104pp; English.
XX
CC The present invention relates to a new human protein kinase polypeptide
CC designated 58848. The invention is useful for modulating 58848-mediated
CC activities which are useful for developing diagnostic and therapeutic
CC agents for protein kinase associated or other 58848-associated disorders
CC such as cellular, proliferative and/or differentiate disorders e.g.
CC cancer, leukaemia; hormonal disorders e.g. diabetes; immune disorders
CC e.g. autoimmune disease; blood vessel disorders e.g. atherosclerosis,
CC hypertension; platelet disorders; cardiovascular disorders e.g. cardiac
CC hypertrophy; heart failure; neurological disorders e.g. ischaemia,
CC Alzheimer's disease, Parkinson's disease, Huntington's disease, acquired
CC immunodeficiency syndrome (AIDS); bone metabolism disorders e.g. rickets,
CC osteoporosis, cirrhosis; haematopoietic neoplastic disorders e.g.
CC Hodgkin's disease, acute leukaemia; liver disorders e.g. Gaucher's
CC disease, viral diseases e.g. Hepatitis B; pain or metabolic disorder

DR WPI; 2002-083100/11.
 XX Forming embryonic germ cells useful as model system to study imprinting
 PT involves mating genetically divergent male and female mammal of same
 PT species, dissecting and dissociating embryo obtained from pregnant
 PT mammal
 XX
 XX Disclosure; Page 115; 125pp; English.
 XX
 CC The present invention relates to a model system for genomic imprinting
 CC using pluripotent mouse embryonic germ (EG) cell lines derived from an
 CC interspecific cross. Also disclosed is a library containing methylated
 CC CpG islands and a method for assaying methylation in one or more
 CC imprintable genes. The gene imprinting assay is carried out by
 CC single-strand conformation polymorphism (SSCP), quantitative sequencing,
 CC carried out to determine the post-translational modification of histones.
 CC The method further involves identifying a test substance as a candidate
 CC drug for treating cancer if the test substance enhances imprinting of a
 CC gene whose imprinting is lost in cancer, or if the test substance
 CC inhibits imprinting of a gene whose imprinting is gained in cancer. The
 CC methylated CpG islands are useful for providing an assessment of the risk
 CC of developing cancer, or for providing diagnostic information relative to
 CC cancer which involves determining the methylation status of the CpG
 CC island in a patient's DNA. The EG cells allow the accession of imprinted
 CC genes which are useful for detecting birth defects, diabetes and cancers
 CC associated with aberrant imprinting. The EG cell lines represent the
 CC first in vitro model system in which genomic imprinting can be followed
 CC dynamically and the two alleles can be distinguished. AAS20970-AAS21029
 CC represent DNA sequences of unknown function given in the present
 CC invention.
 CC Note: The present sequence is given in the Sequence listing but is not
 CC mentioned elsewhere in the patent.
 XX
 XX Sequence 788 BP; 111 A; 291 C; 300 G; 86 T; 0 other;
 Query Match 60.2%; Score 767.6; DB 24; Length 788;
 Best Local Similarity 99.4%; Pred. No. 9e-120;
 Matches 781; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
 QY 420 CCTCTCCCGAGTGGGGCTCCCTGAGGACAGCGTGAAGCGCTGTGTGACAGCTGGGGCT 479
 DB 785 CGCCCGCAGGTGGGGCTCCCTGAGGACAGCGTGAAGCGCTGTGTGACAGCTGGGGCT 726
 QY 480 GCGCTGGACTTCATGCACGCGGCGCAGCTGTGTGACGACATCAAGCCCGAGAACGT 539
 DB 725 GCGCTGGACTTCATGCACGCGGCGCAGCTGTGTGACGACATCAAGCCCGAGAACGT 666
 QY 540 GCTGTGTTCGACCGGAGTGGCGCGGCTAAAGCTGGCGGACTTCGGCATGACGCGCG 599
 DB 665 GCTGTGTTCGACCGGAGTGGCGCGGCTAAAGCTGGCGGACTTCGGCATGACGCGCG 606
 QY 600 CQTGGGCTGCCGCTCAAGCGGTGAGCGGCACCATCCCTTACAGCGGCGCTGAGGTGTG 659
 DB 605 CQTGGGCTGCCGCTCAAGCGGTGAGCGGCACCATCCCTTACAGCGGCGCTGAGGTGTG 546
 QY 660 CCAGCGGCGCGCGCGAGTGGCGGTGACACGCGCGTGGACGTGTGGCGCTTCGG 719
 DB 545 CCAGGCGGCGCGCGCGAGTGGCGGTGACACGCGCGTGGACGTGTGGCGCTTCGG 486
 QY 720 CQTGTCTATCTTCTGGTGTCTACCGGCAACTTCCGCTGGGAGCGGCTGGGCGCGCA 779
 DB 485 CQTGTCTATCTTCTGGTGTCTACCGGCAACTTCCGCTGGGAGCGGCTGGGCGCGCA 427
 QY 780 CCGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGCGCGCTCCCTTCCGCTTCGCA 839
 DB 426 CCGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGGCGCGCTCCCTTCCGCTTCGCA 367
 QY 840 GTGGCGCGCGCTTCACGAGCCCGCGTGGCGCATGTGTTCACGCGCTTACTGGCCCTCGAGCC 899
 DB 366 GTGGCGCGCGCTTCACGAGCCCGCGTGGCGCATGTGTTCACGCGCTTACTGGCCCTCGAGCC 307
 QY 900 CGAGCGCGCGCGCGCGCAAGGAGTGTTCGCTTCTCCTCAAGCAGCGCTCACGTCCGA 959

DB 306 CGAGCGCGCGCGCGCGCAAGGAGTGTTCGCTTCTCAAGCAGCGCTCACGTCCGA 247
 QY 960 GCTGCGCGCGCGCGCGCTTCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCG 1019
 DB 246 GCTGCGCGCGCGCGCGCTTCGACCGCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCG 187
 QY 1020 CGGCGCACGTGCGCTCGAGCGCGCTGGCGCGCTCAAGCGCGCGCGCTGCTGACCGAGCGG 1079
 DB 186 CGGCGCACGTGCGCTCGAGCGCGCTGGCGCGCTCAAGCGCGCGCGCTGCTGACCGAGCGG 127
 QY 1080 CAGCGGCTCCCG 1139
 DB 126 CAGCGGCTCCCG 67
 QY 1140 GGTGCCAGTCCCG 1199
 DB 66 GGTGCCAGTCCCG 7
 QY 1200 CGGCGCG 1205
 DB 6 CGGCGCG 1
 RESULT 13
 AAS21038/c
 ID AAS21038 standard; DNA; 788 BP.
 XX
 XX AAS21038;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Isolated CpG island 2-42.
 XX
 XX Human; genomic imprinting; pluripotent mouse embryonic germ cell line;
 XX EG; methylated CpG island; DNA methylation; gene imprinting;
 XX post-translational modification of histone; cancer; birth defect;
 XX diabetes; aberrant imprinting; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200190313-A2.
 XX
 PD 29-NOV-2001.
 PF 22-MAY-2001; 2001WO-US16253.
 XX
 PR 22-MAY-2000; 2000US-206158P.
 PR 22-MAY-2000; 2000US-206161P.
 XX
 PA (UJJO) UNIV JOHNS HOPKINS.
 XX
 PI Feinberg A, Strichman-Almashanu L, Jiang S;
 XX
 XX WPI; 2002-083100/11.
 XX
 XX Forming embryonic germ cells useful as model system to study imprinting
 XX involves mating genetically divergent male and female mammal of same
 XX species, dissecting and dissociating embryo obtained from pregnant
 XX mammal
 XX
 PS Disclosure; Fig 14; 125pp; English.
 XX
 CC The present invention relates to a model system for genomic imprinting
 CC using pluripotent mouse embryonic germ (EG) cell lines derived from an
 CC interspecific cross. Also disclosed is a library containing methylated
 CC CpG islands and a method for assaying methylation in one or more
 CC imprintable genes. The gene imprinting assay is carried out by
 CC single-strand conformation polymorphism (SSCP), quantitative sequencing,
 CC carried out to determine the post-translational modification of histones.
 CC The method further involves identifying a test substance as a candidate
 CC drug for treating cancer if the test substance enhances imprinting of a

Best Local Similarity 86.1%; Pred. No. 1.4e-152; Matches 1098; Conservative 0; Mismatches 156; Indels 21; Gaps 2;			
Qy	1	ATGAGCGTGGGCTCCACAGACCTGAGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG	60
Db	205	ATGAGCGTGGGCTCCCTGAGCTGAAACCGCTCCACTCCCTGCTGTGTGGGCGGGG	264
Qy	61	ACTGCCCTGGGCTGTGGCGGTGTGCCCTTCTCACTGAAGACATGACGCCCTGACT	120
Db	265	GCCGCCCTGTACCTGTGTGCAAGGTGTGCCCTTCTCAAGAGACATCAAGCCCTGACC	324
Qy	121	CTCCGACACTGGCGCGCAGCGCTCACCAAGACTAGTAAGTCCGGGAGCTGGGC	180
Db	325	CTGGCACACTGGCTGCCAGGACGTCAACCAAGCACTACGAGCTCGTCCGGGAGCTGGGT	384
Qy	181	AAAGGCACCTATGGGAAGGTTGACCTGGTGGTCTCAAGGGCACAGGCACAAAATGGCA	240
Db	385	AAAGGACCTACGGGAAGGTCCGACCTGGTGGCTTACAAGGGCACAGGCACCTAAAATGGCC	444
Qy	241	CTGAAGTTGTGAACAGAGACAAACCAAGCTGAAGAACTTCCTACGGGAGGTGAGCATC	300
Db	445	CTGAATTTGTGAATAAGATGAAGCAAGCTGAAGAACTTCCTGCGGTGAGGTGAGCATC	504
Qy	301	ACCAACAGCTCTCCTCCAGCCCTTCAATCATCAAGGTCTTTGACGTGCTTTTGAGACA	360
Db	505	ACCAACAGCTCTGCTTAGCCCTTCAATCATCAAGGTCTTCGACGTGCTTTTGAGACG	564
Qy	361	GAGACTGCTACTGCTTTGCCAGAGTACGACCTGCTGGGACCTCTTTGACATATC	420
Db	565	GAGAGTCTACTGCTTTTGCTCAGGAGTATGACACCTGCTGGGACCTCTTTGACATATC	624
Qy	421	CCTCCAGGTGGGCTCCCTGAGACACAGGTGAAGCCTGTGTGCACAGCTGGGCTG	480
Db	625	CCTCCTCAGGTGGGCTCCCGAGGACACGCTGAAGCCTGTGTGCACAGCTGGGCTG	684
Qy	481	GCGCTGGACTTCATCAGCGGGCGGAGCTGGTGTGACCGCATCAAGCCCGAGAACTG	540
Db	685	GCACTGACTTCATGATAGCAGGAGCTGGTGTGACCGGCATCAAGCCCGAGAACTG	744
Qy	541	CTGCTGTTTCGACCGGAGTCCGCGCTAAAGCTGGCGGACTTCGGCATGACGCGCGC	600
Db	745	CTGCTGTTTCGACCGGAGTCCGCGCTGAAGCTGGCTGACTTCGGCATGACGCGCGC	804
Qy	601	GTGGCTCCCGCTCAAGCGGTGAGCGGACCAATCCCTTACACGGCGCCTGAGGTGTC	660
Db	805	GTAGCTCCCGTGTGAAGCGAGTAAAGCGGCACTATACCTTACACGGCGCCTGAGGTGTC	864
Qy	661	CAGCGGGCCGCGCGCGGCTGGCGGTGGACACGGGCGTGGAGCTGTGGGCTTCGGC	720
Db	865	CAGGCTGGCGCGCGGATGGCTTCCGGTGGACACGGGCGTGGATGTGTGGCATTCGGC	924
Qy	721	GTGCTCAFTCTTCGCTGCTCACCGGCAACTTCCCGTGGGAGGGCGGTTCGGGCGCGAC	780
Db	925	GTGCTCAFTCTTCGCTGCTCACTGGCAACTTCCCGTGGGAGGCTGCTCAAGTGGCGAT	984
Qy	781	GCCTCTTCGAGGATTCGTGGCTGGACGGGCGCGCTGCGGGGCTGCTTCGGAG	840
Db	985	GCCTCTTCGAGGAAATTTGTGCTGGACGGGCGTGCCTGCCCGGCGCTCCCATCCCG	1044
Qy	841	TGGCGCGCTTCACCGAGCCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTGAGGCC	900
Db	1045	TGGCGAGCTTTACGAGGCTGCTCTACGCATGTTTCCAGGGCTTCTGGCGGTGAGGCT	1104
Qy	901	GAGCGCGCGGCCACCAAGAGGTGTTCCGCTTCTCAAGCAGAGCTCACGTCCGAG	960
Db	1105	GAGCGGCTGGCGCGCCCAAGAGGTCTTTCGCTTCTCAAGCATGAGCTCACATCTGAG	1164
Qy	961	CTGCGCGCGGCGCTTCGACCGCGCGGCAAGCCCCCGGGGACCGCGCGCGCGCC	1020
Db	1165	CTGCGCGCGGCGCATCGACCGCGCGGCAAGAGCCCTGCGGACCGCGCTGCTG	1218
Qy	1021	GGGCACTGCGCTCGAGGCGCTGGCGCTCAAGCGGACGGTCTACCGAGAGCGGC	1080
Db			
Db	1219	GGGCCCCCTGGCCCTTGAAGCTCCAGGGCCACTCAAGCGCACTGTGCTCACGAGAGTGGC	1278
Qy	1081	AGCGCTCCCGCGCGCGCCCGCCCTCGGTCGGTGGTCCCTTGCCTGTCGCGTCCGCTGCGG	1140
Db	1279	AGCGCTCGGGCTTCCACACCGCTCCCGCGCTAGGGCCCT-----GGTACCC	1323
Qy	1141	GTGCACTGCGCGTGGCTGCTGTCGCGAGCCGCGCTAGCTCCCGAGGGGCCCCC	1200
Db	1324	GTGCACTGCGCGTGGCTGCTGTCGCGAGCTACCGCTACCTGAGGCTGGTCTGGCTCCACCGCGACCCCGG	1383
Qy	1201	GGCGGACCGACCGCGCGCGGCAAGAGCAAGGAGGAGGTGGTGGTGGCCACGCGCCATC	1260
Db	1384	GGCAGGACCGACCGCGCGCTGGGACAGAGGAGGAGGTGGTGGTGGCCACGCGCCATC	1443
Qy	1261	GAGATCTGGCTTGA	1275
Db	1444	GAGATCTGGCTTGA	1458
RESULT 10			
AAD49417			
ID	AAD49417 standard; DNA; 936 BP.		
XX	AC AAD49417;		
XX	XX		
DT	24-MAR-2003 (first entry)		
XX	Human serine/threonine protein kinase-like protein (STPK) DNA #2.		
DE	Human; serine/threonine protein kinase-like protein; STPK; diabetes;		
KW	central nervous system; CNS; metabolic disease; urology disorder; COPD;		
KW	chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;		
KW	Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;		
KW	epilepsy; obesity; carcinoma; leukaemia; benign prostatic hyperplasia;		
KW	anorectic; neuroprotective; nootropic; cytostatic; uropathic; gene; ds.		
XX	Homo sapiens.		
OS	WO200283882-A2.		
FN	24-OCT-2002.		
PD	12-APR-2002; 2002WO-EP04080.		
PF	12-APR-2001; 2001US-283189P.		
XX	04-JUN-2001; 2001US-294989P.		
PR	30-JUL-2001; 2001US-308094P.		
PR	08-FEB-2002; 2002US-354574P.		
XX	(FARB) BAYER AG.		
PA	Koehler RH;		
XX	WPI; 2003-093018/08.		
FI	New polynucleotides encoding serine/threonine protein kinase-like		
XX	protein (STPK) and its encoded protein, useful for identifying		
PT	modulators of STPK activity, and in gene therapy for treating e.g.		
PT	cancer, obesity or diabetes		
XX	Claim 1; Page 151-152; 161pp; English.		
PS	The invention relates to human serine/threonine protein kinase-like		
XX	protein (STPK) and its nucleic acid. The STPK sequences are useful for		
CC	identifying test compounds, that may act as agonists or antagonists at		
CC	the receptor site and which can be regulated to provide therapeutic		
CC	effects. The reagent (e.g. antibody or antisense oligonucleotide, which		
CC	decreases the expression of human STPK gene or decreases the levels of		
CC	STPK protein), STPK agonist, STPK protein or expression vector is		
CC	useful for treating a patient with a chronic obstructive pulmonary		
CC	disease (COPD), central nervous system (CNS) disorder, cancer, metabolic		
CC	disease, diabetes, obesity or urology disorders. The vectors comprising		
CC	STPK DNA and reagents are useful for preparing a medicament for		

CC idiopathic hyperglycaemia, haematopoietic disorders such as autoimmune
 CC disease including rheumatoid arthritis, osteoarthritis, multiple
 CC sclerosis, systemic lupus erythematosus, psoriasis, conjunctivitis,
 CC ulcerative colitis, asthma allergic asthma, autoimmune uveitis,
 CC aplastic anaemia, Grave's disease, graft-versus-host disease; liver
 CC disorders e.g. Gaucher's disease, Wilson's disease; and platelet
 CC disorders. Numerous other diseases that can be treated or prevented
 CC are listed in the specification. The present sequence is the cDNA
 CC encoding protein kinase 32374.

XX
 SQ Sequence 2893 BP; 500 A; 989 G; 929 G; 466 T; 9 other;

Query Match 88.8%; Score 1131.8; DB 24; Length 2893;

Best Local Similarity 94.8%; Pred. No. 9.5e-181;

Matches 1205; Conservative 0; Mismatches 57; Indels 9; Gaps 3;

QY 14 GCCCAGAGCCTGAGCCGCCCGCTCCCTGACCTGCTGTGGCCGGGAGCTGCCCTGGGC 73
 DB 44 GCCTTGAAACCGGCCCGGCTCTTGACCTTGCTTGTGGCCGGAAGACTGCCCTGGGC 103
 QY 74 CTGGTGCCG--GTGTGCCCTTCTCACTGAAGACATGACGGCCCTGACTCTCCGCACAC 130
 DB 104 CTGGTCCGGGTGTGCCCTTCTCACTTGAAGACATGACGGCCCTGACTCTCCGCACAC 163
 QY 131 TGG--CCGCAGAGCATCTCAAGACACTAGCAACTAGT---CCGGGAGCTGGGCAAG 184
 DB 164 TTGGCCCGCCAGCGACGTCAACCAAGCACCTACGAAACTAGTCCCGGGAGCTGGGCAAG 223
 QY 185 GCACCTATGGAGTTGACTGTGTCTACAGGGCAGAGGCACAGCAAAATGCACTGA 244
 DB 224 GCACCTACGGAGGTGTATCTGGTGTCTACAGGGCAGAGGCACAGCAAAATGCACTGA 283
 QY 245 AGTTTGTGAACAGACAGCAAAACCAAGCTGAAGAACTTCTTACGGGAGGTGAGCATACCA 304
 DB 284 AGTTTGTGAACAGACAGCAAAACCAAGCTGAAGAACTTCTTACGGGAGGTGAGCATACCA 343
 QY 305 ACAGCTCTCTCAGGCCCTTATCATCAAGGTCTTTGACGTGGTCTTTGAGACAGAGG 364
 DB 344 ACAGCTCTCTCAGGCCCTTATCATCAAGGTCTTTGACGTGGTCTTTGAGACAGAGG 403
 QY 365 ACTGCTACGCTTTGCCCGGAGTACGACCTGCTGGGGACCTTTGACATCATCTCCTC 424
 DB 404 ACTGCTACGCTTTGCCCGGAGTACGACCTGCTGGGGACCTTTGACATCATCTCCTC 463
 QY 425 CCCAGGTGGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGCTGGCGC 484
 DB 464 CCCAGGTGGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGCTGGCGC 523
 QY 485 TGGACTTATGACAGGGGGGAGCTGGTGACCCGGACATCAAGCCCGAGACGTGTGC 544
 DB 524 TGGACTTATGACAGGGGGGAGCTGGTGACCCGGACATCAAGCCCGAGAACGTGTGC 583
 QY 545 TGTTTCGACCGGAGTGGCCCGCGTAAAGCTTGGCGACTTCGGCATGACGGCCCGCTGG 604
 DB 584 TGTTTCGACCGGAGTGGCCCGCGTAAAGCTTGGCGACTTCGGCATGACGGCCCGCTGG 643
 QY 605 GCTCCCGCTCAAGCGCTGAGCGGACCATCCCTTACACGGCGCTGAGTGTGCCAGG 664
 DB 644 GCTCCCGCTCAAGCGCTGAGCGGACCATCCCTTACACGGCGCTGAGTGTGCCAGG 703
 QY 665 CGGGCCGCGCAGCGGCTGGGTGGACAGGGGTGGAAGTGTGGGCTTCCGCTGGC 724
 DB 704 CGGGCCGCGCAGCGGCTGGGTGGACAGGGGTGGAAGTGTGGGCTTCCGCTGGC 763
 QY 725 TCATCTTCTGCTGCTACCGGCACTTCCGCTGGGAGCGCGCTGGCGCGCGACGCT 784
 DB 764 TCATCTTCTGCTGCTACCGGCACTTCCGCTGGGAGCGCGCTGGCGCGCGACGCT 823
 QY 785 TCTTCAGGAGTTCGTGCTGGCAGCGGGGCGGCTCCCGGGGTGCTCTTCGAGTGGC 844
 DB 824 TCTTCAGGAGTTCGTGCTGGCAGCGGGGCGGCTCCCGGGGTGCTCTTCGAGTGGC 883
 QY 845 GCCGCTTACCGAGCGCGCTGGCGCATGTTCCAGGCTTACTGCGCTTGGAGCCCGGAGC 904

DB 884 GCCGCTTACCGAGCCCGCTGCGCATGTTCCAGCGCTTACTGCGCTTGGAGCCCGAGC 943
 QY 905 GCGCGGGCCAGCCCAAGGAGGTGTTCCCGTTTCTCAAGCACGAGTCACTCCGAGCTGC 964
 DB 944 GCGCGGGCCAGCCCAAGGAGGTGTTCCCGTTTCTCAAGCACGAGTCACTCCGAGCTGC 1003
 QY 965 GCGCGGGCCCTTGCACACCGCGCGCAAGCCCGCGGGGAGCCCGCGCGCGCGCGGC 1024
 DB 1004 GCGCGGGCCCTTGCACACCGCGCGCAAGCCCGCGGGGAGCCCGCGCGCGCGCGGC 1063
 QY 1025 CACTGCGCTCGAGCGCGCTGCGCGCTCAAGGGACGCTGCTGACCGAGAGCGGCGAGCG 1084
 DB 1064 CACTGCGCTCGAGCGCGCTGCGCGCTCAAGGGACGCTGCTGACCGAGAGCGGCGAGCG 1123
 QY 1085 GTTCCCGCGCGCGCGCGCGCGCTGCGCGCTGCGCGCTTGCCTTGCCTGCGCGTGC 1144
 DB 1124 GCTCCCGCGCGCGCGCGCGCGCTGCGCGCTGCGCGCTTGCCTTGCCTGCGCGTGC 1183
 QY 1145 CAGTGCCTGCGCGGTGCTGCGCGAGCCCGCGCTAGCTCCCGAGGGGCGCGCGCGGC 1204
 DB 1184 CAGTGCCTGCGCGGTGCTGCGCGAGCCCGCGCTAGCTCCCGAGGGGCGCGCGCGGC 1243
 QY 1205 GGAACCGCGCGCGCGCGCGAGCAAGAGCAAGGGCAGGTGCTGCGCACCGCCATCGAGA 1264
 DB 1244 GGAACCGCGCGCGCGCGAGCAAGAGCAAGGGCAGGTGCTGCGCACCGCCATCGAGA 1303
 QY 1265 TCTGCGTCTGA 1275
 DB 1304 TCTGCGTCTGA 1314

RESULT 9

AAA09825

ID AAA09825 standard; DNA; 1527 BP.

XX AAA09825;

XX 05-JUL-2000 (first entry)

XX Rat PKs gene.

XX PKs protein; Wistar rat; cerebral nervous system disease;

XX neurological function-related disease; ds.

XX Rattus sp.

XX JP2000060571-A.

XX 29-FEB-2000.

XX 20-AUG-1998; 98JP-0249064.

XX 20-AUG-1998; 98JP-0249064.

XX (MITU) MITSUBISHI CHEM CORP.

XX WPI; 2000-249682/22.

XX P-PSDB; AAY83904.

XX Novel mammalian peptide and a polynucleotide encoding it - useful for
 PT treatment and diagnosis of cerebral nervous system diseases and
 PT neurological function-related diseases
 XX Claim 15; Page 10-11; 15pp; Japanese.

XX This sequence represents the gene encoding the PKs protein from Wistar
 CC rats. The peptide is useful for treatment and diagnosis of cerebral
 CC nervous system diseases and neurological function-related diseases.

XX Sequence 1527 BP; 263 A; 495 C; 501 G; 268 T; 0 other;

Query Match 75.6%; Score 953.4; DB 21; Length 1527;

181 AAAGGACCTATGGAAGGTTGACCTGTGTCTACAAAGGACACAGGCACAAAAATGSCA 240
181 AAAGGACCTATGGAAGGTTGACCTGTGTCTACAAAGGACACAGGCACAAAAATGSCA 240
241 CTGAAGTTTGAACAGAGCAAAACCAAGCTGAGACTTCCCTACGGGAGGTGAGCATC 300
241 CTGAAGTTTGAACAGAGCAAAACCAAGCTGAGACTTCCCTACGGGAGGTGAGCATC 300
301 ACCAACAGCCCTCTCTCCAGGCCCTTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA 360
301 ACCAACAGCCCTCTCTCCAGGCCCTTTCATCATCAAGGTCTTTGACGTGGTCTTTGAGACA 360
361 GAGACTGCTACGCTTTTGGCCAGGAGTACGACCTGCTGGGGACCTTTTGACATCATC 420
361 GAGACTGCTACGCTTTTGGCCAGGAGTACGACCTGCTGGGGACCTTTTGACATCATC 420
421 CCTCCAGGTGGGCTCCCTGAGGACACGCTGAGCGCTGTGTGACAGACTGGGCTG 480
421 CCTCCAGGTGGGCTCCCTGAGGACACGCTGAGCGCTGTGTGACAGACTGGGCTG 480
481 GCGTGTGACTTCATGACGGGCGGAGCTGTGTGACCGCGACATCAAGCCCGAGAACCTG 540
481 GCGTGTGACTTCATGACGGGCGGAGCTGTGTGACCGCGACATCAAGCCCGAGAACCTG 540
541 CTGCTGTTCGACCGGAGTGGCCGCGCTAAAGCTGGCGCGACTTCGGCATGACGCGCGCG 600
541 CTGCTGTTCGACCGGAGTGGCCGCGCTAAAGCTGGCGCGACTTCGGCATGACGCGCGCG 600
601 GTGGCTCCCGCTCAAGCGGTGAGCGGACCATCTTACAGCGGCGCTGAGGTGTC 660
601 GTGGCTCCCGCTCAAGCGGTGAGCGGACCATCTTACAGCGGCGCTGAGGTGTC 660
661 CAGCGGCGCGCGCGCTGCGGTGACACACGCGCGCTGAGCGCTGTGGGCTTCGCG 720
661 CAGCGGCGCGCGCGCTGCGGTGACACACGCGCGCTGAGCGCTGTGGGCTTCGCG 720
721 GTGCTCATCTTCTGCTGTCTACCGGCACTTCCGCTGGGAGGCGGCTGCGGCGCCGAC 780
721 GTGCTCATCTTCTGCTGTCTACCGGCACTTCCGCTGGGAGGCGGCTGCGGCGCCGAC 780
781 GCCTTCTTCGAGGAGTCTGCTGCTGCGGAGCGGCGGCGCTGCGGCGCTGCGGAG 840
781 GCCTTCTTCGAGGAGTCTGCTGCTGCGGAGCGGCGGCGCTGCGGCGCTGCGGAG 840
841 TGGCGCGCTTCACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTGAGGCGC 900
841 TGGCGCGCTTCACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTGAGGCGC 900
901 GAGCGCGCGGCGGCTCAAGAGGTGTTCGCTTCTCAAGCAGAGCTACGCTCCGAG 960
901 GAGCGCGCGGCGGCTCAAGAGGTGTTCGCTTCTCAAGCAGAGCTACGCTCCGAG 960
961 CTGCGCGCGGCGGCTTCGACCGCGCGCAAGCGGCGGCGGCGGCGGCGGCGGCGG 1020
961 CTGCGCGCGGCGGCTTCGACCGCGCGCAAGCGGCGGCGGCGGCGGCGGCGGCGG 1020
1021 GGGCCACTCGGCTCGAGGCGCTGGGCGCTCAAGCGGAGCGGTGCTACCGAGAGCGCG 1080
1021 GGGCCACTCGGCTCGAGGCGCTGGGCGCTCAAGCGGAGCGGTGCTACCGAGAGCGCG 1080
1081 AGCGGCTCCGCGCGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 1140
1081 AGCGGCTCCGCGCGGCGGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 1140
1141 GTGCGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1141 GTGCGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
1201 GSCCGGACCGAGCGGCGGCGGACAAAGGCGAGGTGCTGCTGCGCACGCGCCATC 1260
1201 GSCCGGACCGAGCGGCGGCGGACAAAGGCGAGGTGCTGCTGCGCACGCGCCATC 1260

Qy 1261 GAGATCTGCGTC 1272
Db 1261 GAGATCTGCGTC 1272
RESULT 8
ABK10101
ID ABK10101 standard; cDNA; 2893 BP.
XX ABK10101;
AC ABK10101;
XX 21-MAY-2002 (first entry)
XX Human cDNA encoding protein kinase 32374.
XX Human; ss; gene; protein kinase; 32374; 18431; cancer; infection;
cellular proliferation; pain; inflammation; metabolic disorder; obesity;
lipid disorder; diabetes; brain disorder; Alzheimer's disease;
neural tube defect; cerebrovascular disease; hypoxia; AIDS;
acquired immunodeficiency syndrome associated myopathy; heart disorder;
multiple sclerosis; angina pectoris; blood vessel disorder; vasculitis;
bone disorder; osteoporosis; haematopoietic disorder; autoimmune disease;
rheumatoid arthritis; osteoarthritis; multiple sclerosis;
systemic lupus erythematosus; asthma; Grave's disease; liver disorder;
graft-versus-host disease.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
CDS 274..1314
FT /**tag= a
FT /product= "Protein kinase 32374"
XX
PN WO200210401-A2.
XX
PD 07-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-US23653.
XX
PR 28-JUL-2000; 2000US-221543P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA Meyers R, Kapeller-Libermann R, Silos-Santiago I;
XX WPI; 2002-188748/24.
XX P-PSDB; AAU76323.
XX
XX Novel protein kinase nucleic acid molecules and the encoded proteins
for diagnosing and treating cellular proliferative, bone, immune,
cardiovascular, liver, pain or metabolic disorders and identifying
modulators -
XX
XX Claim 1; Fig 1; 141pp; English.
XX
XX The invention relates to an isolated protein kinase polypeptide termed as
32374 or 18431, their allelic variants, fragments and nucleic acids
encoding them. Also included are a host cell containing the nucleic
acid, an antibody that binds to the protein kinase and modulators
of the activity of the kinase. The proteins, nucleic acids,
antibodies and modulators are useful for diagnosing, treating or
monitoring cancer or aberrant cellular proliferation and/or
differentiation such as lung, ovarian or brain cancer, pain (e.g.
inflammation or infection) or metabolic disorder (e.g. obesity, anorexia
disease), neural tube defects, cerebrovascular diseases (e.g. hypoxia,
ischaemia, infarction), AIDS (acquired immunodeficiency syndrome) -
associated myopathy, multiple sclerosis, disorders of the heart
(e.g. heart failure, angina pectoris, myocardial infarction),
blood vessel disorders (e.g. arteriovenous fistula, vasculitis,
syphilitic aneurysms), disorders of the bone (e.g. osteoporosis,
cirrhosis, hypoparathyroidism, hyperparathyroidism, tropical sprue,

```
QY 421 CTTCCACAGTGGGCTCCCTGAGGACACCGGTGAAGCGTGTGTGACGAGCTGGGCGCTG 480
Db 777 CTTCCACAGTGGGCTCCCTGAGGACACCGGTGAAGCGTGTGTGACGAGCTGGGCGCTG 836
QY 481 GCGGTGACATTCATGACGGGCGGCGAGCTGCTGTCACCGGACATCAAGCCCGAGAGCGTG 540
Db 837 GCGGTGACATTCATGACGGGCGGCGAGCTGCTGTCACCGGACATCAAGCCCGAGAGCGTG 896
QY 541 CTGTGTTGACCGCGAGTGCCTGCGGCTAAAGCTGGCCGACTTGGSCATGACGCGCGCG 600
Db 897 CTGTGTTGACCGCGAGTGCCTGCGGCTAAAGCTGGCCGACTTGGSCATGACGCGCGCG 956
QY 601 GTGGGCTGCGGCTAAAGCGGCTGAGCGGACCATCCCTTACAGGGGCGCTGAGGTGTGC 660
Db 957 GTGGGCTGCGGCTAAAGCGGCTGAGCGGACCATCCCTTACAGGGGCGCTGAGGTGTGC 1016
QY 661 CAGGCGGCGCGCGCGGCTGGCGGTGACACGGGCGTGGACGCTGTGGGCTTCGCGC 720
Db 1017 CAGGCGGCGCGCGCGGCTGGCGGTGACACGGGCGTGGACGCTGTGGGCTTCGCGC 1076
QY 721 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCCGTGGAGGCGGCTCGGGCGCGAC 780
Db 1077 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCCGTGGAGGCGGCTCGGGCGCGAC 1136
QY 781 GCCTTCTTCAGGAGTTCTGTCGCTGTCAGCGGGGCGGCTGCGGGGCTGCCCTTCGAG 840
Db 1137 GCCTTCTTCAGGAGTTCTGTCGCTGTCAGCGGGGCGGCTGCGGGGCTGCCCTTCGAG 1196
QY 841 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTTGGCCCTGGAGCCC 900
Db 1197 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCCAGCGCTTACTTGGCCCTGGAGCCC 1256
QY 901 GAGCGCGCGCGCGCGCGGCTGCGCTTCCCTCAAGCAGCAGCTCAAGTCCGAG 960
Db 1257 GAGCGCGCGCGCGCGGCTGCGCTTCCCTCAAGCAGCAGCTCAAGTCCGAG 1316
QY 961 CTGGCGCGCGGCGCTGCGACCGCGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 1317 CTGGCGCGCGGCGCTGCGACCGCGCGGCGAGCGCGCGCGCGCGCGCGCGCGCG 1376
QY 1021 GGGCCACTGCGCTCGAGGCGCTGGCGCGCTCAAGCGGACGGTGTGACGAGAGCGCG 1080
Db 1377 GGGCCACTGCGCTCGAGGCGCTGGCGCGCTCAAGCGGACGGTGTGACGAGAGCGCG 1436
QY 1081 AGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db 1437 GCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1496
QY 1141 GTGCCAGTCCCGTCCCGTCCCTGTGCGCGAGCGCGCGCTAGCTCCCGAGGCGCGCG 1200
Db 1497 GTGCCAGTCCCGTCCCGTCCCTGTGCGCGAGCGCGCGCTAGCTCCCGAGGCGCGCG 1556
QY 1201 GCGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
Db 1557 GCGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1616
QY 1261 GAGATCTGCGTCTGA 1275
Db 1617 GAGATCTGCGTCTGA 1631
```

RESULT 7

```
AAD49416
ID AAD49416 standard; DNA; 1272 BP.
XX
AC AAD49416;
XX
XX
DT 24-MAR-2003 (first entry)
DE Human serine/threonine protein kinase-like protein (STPKP) DNA #1.
DE Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
XX
```

Chronic obstructive pulmonary disease; multiple sclerosis; gene therapy; Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer; epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia; anorectic; neuroprotective; nootropic; cytostatic; uropathic; gene; ds.

Homo sapiens.

Key Location/Qualifiers
CDS 1..1272
FT /tag= a
FT /product= "Human STPKP protein"
FT /note= "CDS does not include stop codon"
FT /partial

W0200283882-A2.

24-OCT-2002.

12-APR-2002; 2002NO-BP04080.

12-APR-2001; 2001US-283189P.

04-JUN-2001; 2001US-294989P.

30-JUL-2001; 2001US-308094P.

08-FEB-2002; 2002US-354574P.

(FARB) BAYER AG.

Koehler RH;

WPI; 2003-093018/08.

P-PSDB; AAE32010.

New polynucleotides encoding serine/threonine protein kinase-like protein (STPKP) and its encoded protein, useful for identifying modulators of STPKP activity, and in gene therapy for treating e.g. cancer, obesity or diabetes.

Claim 1; Page 146-147; 161pp; English.

The invention relates to human serine/threonine protein kinase-like protein (STPKP) and its nucleic acid. The STPKP sequences are useful for identifying test compounds, that may act as agonists or antagonists at the receptor site and which can be regulated to provide therapeutic effects. The reagent (e.g. antibody or antisense oligonucleotide, which decreases the expression of human STPKP gene or decreases the levels of STPKP protein), STPKP agonist, STPKP protein or expression vector is useful for treating a patient with a chronic obstructive pulmonary disease (COPD), central nervous system (CNS) disorder, cancer, metabolic disease, diabetes, obesity or urology disorders. The vectors comprising STPKP DNA and reagents are useful for preparing a medicament for modulating the activity of STPKP in the diseases. These diseases include multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy, urinary incontinence, carcinoma, leukaemia, or benign prostatic hypoplasia. STPKP gene is used in gene therapy. The present sequence is human STPKP DNA.

Sequence 1272 BP; 191 A; 450 C; 430 G; 201 T; 0 other;

Query Match 99.0%; Score 1262.4; DB 25; Length 1272;

Best Local Similarity 99.5%; Pred. No. 1.3e-202;

Matches 1266; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGCGCGGG 60

Db 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGCGCGGG 60

QY 61 ACTGCCCTGGGCTGGTGCCTTCTCACTGAACACATGACGAGCCCTGACT 120

Db 61 ACTGCCCTGGGCTGGTGCCTTCTCACTGAACACATGACGAGCCCTGACT 120

QY 121 CTCGGCACACTGGCGCGCGAGCGACGTCACCAAGCACTACGAAGTGTCCGGGAGCTGGGC 180

Db 121 CTCGGCACACTGGCGCGCGAGCGACGTCACCAAGCACTACGAAGTGTCCGGGAGCTGGGC 180

```
Db 985 CTGCTGTTGACCCGCGAGTCCGCGCGGTAAAGCTGGCCGACTTCGGATGACCGCGCGC 1044
Qy 601 GTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACCACTCCCTTACACGGCGCTTGAAGTGTGC 660
Db 1045 GTGGGCTGCGCGCTCAAGCGCGTGAAGCGGACCACTCCCTTACACGGCGCTTGAAGTGTGC 1104
Qy 661 CAGGGGGCGCGCGCGCGCTGCGGTGACACGGGCGTGAAGTGTGCGGCTTGGC 720
Db 1105 CAGGGGGCGCGCGCGCGCTGCGGTGACACGGGCGTGAAGTGTGCGGCTTGGC 1164
Qy 721 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGCTGGGAGGCGGCTCGGGCGCGAC 780
Db 1165 GTGCTCATCTTCTGCTGCTCACCGGCAACTTCCGCTGGGAGGCGGCTCGGGCGCGAC 1224
Qy 781 GCCTTCTCGAGAGTGTGCTGCTGCGTGAAGCGGCGCGCTTGAAGTGTGCGGCTTGGC 840
Db 1225 GCCTTCTCGAGAGTGTGCTGCTGCGTGAAGCGGCGCGCTTGAAGTGTGCGGCTTGGC 1284
Qy 841 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCAGCGCTTACTGGCCCTGGAGCC 900
Db 1285 TGGCGCGCTTACCGAGCGCGCTGCGCATGTTCAGCGCTTACTGGCCCTGGAGCC 1344
Qy 901 GAGCGCGCGCGCGCGCGAGTGTTCGCTTCTTCAAGCAGAGCTCACGTCGCGAG 960
Db 1345 GAGCGCGCGCGCGCGCGAGTGTTCGCTTCTTCAAGCAGAGCTCACGTCGCGAG 1404
Qy 961 CTGGCGCGCGCGCTTGCACCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 1405 CTGGCGCGCGCGCTTGCACCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCG 1464
Qy 1021 GGGCACTCGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGAGCGTGTGACCGAGCGCG 1080
Db 1465 GGGCACTCGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGAGCGTGTGACCGAGCGCG 1524
Qy 1081 AGCGGCTTCCGCGCGCGCGCGCGCGCTGCGGCTGCGTGGCGTGGCGTGGCGTGGCG 1140
Db 1525 GCGGCTTCCGCGCGCGCGCGCGCGCTGCGGCTGCGTGGCGTGGCGTGGCGTGGCG 1584
Qy 1141 GTGCGAGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 1200
Db 1585 GTGCGAGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 1644
Qy 1201 GCGCGGACCGCGCGCGCGCGCAAGAGCAAGAGCGAGTGTGCTGGCCACGCGCATC 1260
Db 1645 GCGCGGACCGCGCGCGCGCGCAAGAGCAAGAGCGAGTGTGCTGGCCACGCGCATC 1704
Qy 1261 GAGATCTGCGTCTGA 1275
Db 1705 GAGATCTGCGTCTGA 1719

RESULT 6
AAD49420
ID AAD49420 standard; DNA; 3124 BP.
XX
AC AAD49420;
XX
DT 24-MAR-2003 (first entry)
XX
DE Human serine/threonine protein kinase-like protein (STPKP) DNA #5.
XX
KW Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia;
KW anorectic; neuroprotective; nootropic; cytostatic; uropathic; gene; ds.
XX
OS Homo sapiens.
XX
FN WO200283882-A2.
XX
```

```
PD 24-OCT-2002.
XX
PF 12-APR-2002; 2002WO-EP04080.
XX
PR 12-APR-2001; 2001US-283189P.
PR 04-JUN-2001; 2001US-294989P.
PR 30-JUL-2001; 2001US-308094P.
PR 08-FEB-2002; 2002US-354574P.
XX
PA (FARB ) BAYER AG.
XX
XX Koehler RH;
XX
XX WPI; 2003-093018/08.
XX
XX New polynucleotides encoding serine/threonine protein kinase-like
XX protein (STPKP) and its encoded protein, useful for identifying
XX modulators of STPKP activity, and in gene therapy for treating e.g.
XX cancer, obesity or diabetes .
XX
XX Claim 1; Page 159-161; 161pp; English.
XX
XX The invention relates to human serine/threonine protein kinase-like
XX protein (STPKP) and its nucleic acid. The STPKP sequences are useful for
XX identifying test compounds, that may act as agonists or antagonists at
XX the receptor site and which can be regulated to provide therapeutic
XX effects. The reagent (e.g. antibody or antisense oligonucleotide, which
XX decreases the expression of human STPKP gene or decreases the levels of
XX STPKP protein), STPKP agonist, STPKP protein or expression vector is
XX useful for treating a patient with a chronic obstructive pulmonary
XX disease (COPD), central nervous system (CNS) disorder, cancer, metabolic
XX disease, diabetes, obesity or urology disorders. The vectors comprising
XX STPKP DNA and reagents are useful for preparing a medicament for
XX modulating the activity of STPKP in the diseases. These diseases include
XX multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
XX urinary incontinence, carcinoma, leukaemia, or benign prostatic
XX hypoplasia. STPKP gene is used in gene therapy. The present sequence is
XX human STPKP DNA.
XX
XX Sequence 3124 BP; 485 A; 1110 C; 1081 G; 448 T; 0 other;
XX
XX Query Match 99.9%; Score 1273.4; DB 25; Length 3124;
XX Best Local Similarity 99.9%; Pred. No. 2e-204;
XX Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAGCGTGGCTGCGCGCGAGCGCTGAGCGCGCGCGCTGAGCGCTGCTGCGCGCGCGG 60
Db 357 ATGAGCGTGGCTGCGCGCGAGCGCTGAGCGCGCGCGCTGAGCGCTGCTGCGCGCGCGG 416
Qy 61 ACTGCGCGCTGGCGCTGCGCGGTGCGCGCTTCTCACTGAAGACATGCGAGCGCTGACT 120
Db 417 ACTGCGCGCTGGCGCTGCGCGGTGCGCGCTTCTCACTGAAGACATGCGAGCGCTGACT 476
Qy 121 CTCGCGCACCTGGCGCGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 477 CTCGCGCACCTGGCGCGCGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 536
Qy 181 AAGGCGACCTATGGGAAGGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 537 AAGGCGACCTATGGGAAGGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 596
Qy 241 CTGAAGTGTGACAAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 300
Db 597 CTGAAGTGTGACAAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 656
Qy 301 ACCAACAGCTCTCTCTCCAGCGCGCTTCTCATCAAGTCTTTCAGCTGCTGCTTTCAGACA 360
Db 657 ACCAACAGCTCTCTCTCCAGCGCGCTTCTCATCAAGTCTTTCAGCTGCTGCTTTCAGACA 716
Qy 361 GAGGACTGTGCTGCTTTCGCGGAGTACGACCTGCTGCGGACCTGTTTGACATCATC 420
Db 717 GAGGACTGTGCTGCTTTCGCGGAGTACGACCTGCTGCGGACCTGTTTGACATCATC 776
```


Db 1021 GGGCCACTGCGCTCGAGGCGCCTGGCGCGCTCAAGCGACGGTGTGACCGAGAGCGCG 1080
 Qy 1081 AGCGGCTCCCGGCGCGCGCGCGCGCGCTCGGTCGCGTCCCTTGCCTGCGCGTCCGCG 1140
 Db 1081 GGGGCTCCCGGCGCGCGCGCGCGCGCTCGGTCGCGTCCCTTGCCTGCGCGTCCGCG 1140
 Qy 1141 GTGCCAGTCCCGGCGCGCGTCCCTGTGCGCGAGCGCGCTAGCTCCCGAGGCGCGCCC 1200
 Db 1141 GTGCCAGTCCCGGCGCGCGTCCCTGTGCGCGAGCGCGCTAGCTCCCGAGGCGCGCCC 1200
 Qy 1201 GCGCGACCGAGCGCGCGCGCGACAGAGCAAGAGCAAGAGCGAGGTGTGCTGGCCACGSCCATC 1260
 Db 1201 GCGCGACCGAGCGCGCGCGCGACAGAGCAAGAGCGAGGTGTGCTGGCCACGSCCATC 1260
 Qy 1261 GAGATCTGCGTCTGA 1275
 Db 1261 GAGATCTGCGTCTGA 1275

RESULT 4

AAD49419
 ID AAD49419 standard; DNA; 1275 BP.
 AC AAD49419;
 DT 24-MAR-2003 (first entry)
 DE Human serine/threonine protein kinase-like protein (STPKP) DNA #4.
 KW Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
 KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
 KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
 KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
 KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia;
 KW anorectic; neuroprotective; nootropic; cytostatic; uropathic; gene; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1275
 FT /tag= a
 FT /product= "Human STPKP protein"

WO200283882-A2.

24-OCT-2002.

12-APR-2002; 2002WO-EP04080.

12-APR-2001; 2001US-283189P.

04-JUN-2001; 2001US-294989P.

30-JUL-2001; 2001US-308094P.

08-FEB-2002; 2002US-354574P.

(FARB) BAYER AG.

Koehler RH;

WPI; 2003-093018/08.

P-PSDB; AAE32011.

PT New polynucleotides encoding serine/threonine protein kinase-like
 PT protein (STPKP) and its encoded protein, useful for identifying
 PT modulators of STPKP activity, and in gene therapy for treating e.g.
 PT cancer, obesity or diabetes -

Claim 1; Page 156-157; 161pp; English.

CC The invention relates to human serine/threonine protein kinase-like
 CC protein (STPKP) and its nucleic acid. The STPKP sequences are useful for
 CC identifying test compounds, that may act as agonists or antagonists at
 CC the receptor site and which can be regulated to provide therapeutic
 CC effects. The reagent (e.g. antibody or antisense oligonucleotide, which

CC decreases the expression of human STPKP gene or decreases the levels of
 CC STPKP protein), STPKP agonist, STPKP protein or expression vector is
 CC useful for treating a patient with a chronic obstructive pulmonary
 CC disease (COPD), central nervous system (CNS) disorder, cancer, metabolic
 CC disease, diabetes, obesity or urology disorders. The vectors comprising
 CC STPKP DNA and reagents are useful for preparing a medicament for
 CC modulating the activity of STPKP in the diseases. These diseases include
 CC multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy,
 CC urinary incontinence, carcinoma, leukaemia, or benign prostatic
 CC hypoplasia. STPKP gene is used in gene therapy. The present sequence is
 CC human STPKP DNA.

XX
 SQ Sequence 1275 BP; 191 A; 450 C; 434 G; 200 T; 0 other;

Query Match 99.9%; Score 1273.4; DB 25; Length 1275;

Best Local Similarity 99.9%; Pred. No. 1.9e-204;

Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCTCCCTGACCTCTGTGGCCCGGG 60
 Db 1 ATGAGCGTGGGCTGCCAGAGCCTGAGCGCGCGCTCCCTGACCTCTGTGGCCCGGG 60
 Qy 61 ACTGCCCTGGGCTGGTGGCGGTGCGCTTCTCACTGAAGACATGCAGGCCCTGACT 120
 Db 61 ACTGCCCTGGGCTGGTGGCGGTGCGCTTCTCACTGAAGACATGCAGGCCCTGACT 120
 Qy 121 CTCGCACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGAGCTGGC 180
 Db 121 CTCGCACTGGCGCGCAGCGAGCTCACCAAGCACTACGAACCTAGTCCGGAGCTGGC 180
 Qy 181 AAGGCACTTGGGAAGGTGACCTGGTGTACAAAGGCACAGCACAAATGGCA 240
 Db 181 AAGGCACTTGGGAAGGTGACCTGGTGTACAAAGGCACAGCACAAATGGCA 240
 Qy 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 300
 Db 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 300
 Qy 301 ACCAAGCCTCTCTCCAGCGCCCTTCATCATCAAGGTCTTTGACGTGTTTGAGACA 360
 Db 301 ACCAAGCCTCTCTCCAGCGCCCTTCATCATCAAGGTCTTTGACGTGTTTGAGACA 360
 Qy 361 GAGGACTGTACGTCTTTGCGCAGGAGTACGACCTGCTGGGACCTGTTTGACATATC 420
 Db 361 GAGGACTGTACGTCTTTGCGCAGGAGTACGACCTGCTGGGACCTGTTTGACATATC 420
 Qy 421 CCTCCCAAGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACGAGCTGGGCTG 480
 Db 421 CCTCCCAAGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTGACGAGCTGGGCTG 480
 Qy 481 GCGTGTGACTTTCATGCAAGCGCGGAGTGGTGTGACCGGACATCAAGCCGAGAACTG 540
 Db 481 GCGTGTGACTTTCATGCAAGCGCGGAGTGGTGTGACCGGACATCAAGCCGAGAACTG 540
 Qy 541 CTGCTGTTCGACCGGAGTGGCGCGGTAAAGCTGGCGGACTTCGCGCATGACCGCGCG 600
 Db 541 CTGCTGTTCGACCGGAGTGGCGCGGTAAAGCTGGCGGACTTCGCGCATGACCGCGCG 600
 Qy 601 GTGGGCTGCCCGCTCAAGCGCGTGAAGCGGACCATCCCTTACCGCGCTGAGGTGTGC 660
 Db 601 GTGGGCTGCCCGCTCAAGCGCGTGAAGCGGACCATCCCTTACCGCGCTGAGGTGTGC 660
 Qy 661 CAGCGGGCGCGCGGCTGGCGTGGACACCGGCGGTGGACACGCGGCTTGGGCTTCCGG 720
 Db 661 CAGCGGGCGCGCGGCTGGCGTGGACACCGGCGGTGGACACGCGGCTTGGGCTTCCGG 720
 Qy 721 GTGCTCATCTTCGTGCTCACCGGCAACTTCCGTGGAGGCGGCTGGGCGCGCGAC 780
 Db 721 GTGCTCATCTTCGTGCTCACCGGCAACTTCCGTGGAGGCGGCTGGGCGCGCGAC 780
 Qy 781 GCCTTCTTCAGAGGATTTGTCGCTGCTGCGCTGCGAGCGGGCGGCTTCCGTTCGCG 840
 Db 781 GCCTTCTTCAGAGGATTTGTCGCTGCTGCGCTGCGAGCGGGCGGCTTCCGTTCGCG 840

XX Walke DW, Maricar M, Yu X, Friddle CJ;
XX WPI; 2002-527921/56.
XX P-PSDB; AAU79585.
XX Novel nucleic acid molecule encoding a human kinase, useful in
PT therapeutic, diagnostic and pharmacogenomic applications, as DNA
PT markers for restriction fragment length polymorphism analysis and in
PT forensic biology -
XX Disclosure; Page 37; 37pp; English.
PS The invention discloses an isolated human kinase polypeptide, designated
CC novel human protein (NHP), and the polynucleotide encoding it. Kinases
CC mediate phosphorylation of a wide variety of proteins and compounds in
CC the cell and are also involved in a range of regulatory pathways. The
CC polynucleotide is useful in therapeutic, diagnostic and pharmacogenomic
CC applications, and for identifying compounds that modulate, i.e. act as
CC agonists or antagonists of, the gene expression or gene product
CC activity. The polynucleotide is also useful as a probe in microarrays
CC or other cloning and/or assay formats, for screening collections of
CC genetic material from patients who have a particular medical condition,
CC for identifying mutations associated with a particular disease and also
CC as a diagnostic or prognostic assay. It can also be useful for the
CC detection of mutant human proteins, or inappropriately expressed
CC proteins, for the diagnosis of disease, for screening for drugs effective
CC in perturbing the normal function of the protein, for generation of
CC antibodies and as reagents in assays for screening for compounds that can
CC be used as pharmaceutical agents in the therapeutic treatment of mental,
CC biological or medical disorders and diseases. The sequence presented is
CC the novel human protein (NHP) cDNA which was isolated from human brain
CC and skeletal muscle cDNA libraries.
XX Sequence 1473 BP; 215 A; 533 C; 509 G; 216 T; 0 other;
SQ
Query Match 100.0%; Score 1275; DB 24; Length 1473;
Best Local Similarity 100.0%; Pred. No. 1e-204;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGCGTGGGTGCGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 60
DB 117 ATGAGCGTGGGTGCGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGGCGGGG 176
QY 61 ACTGCCCTTGGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
DB 177 ACTGCCCTTGGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 236
QY 121 CTCGCCACACTGGCCGCCAGCGAGCTCAGAGCACTACGAACTAGTCCGGAGCTGGGC 180
DB 237 CTCGCCACACTGGCCGCCAGCGAGCTCAGAGCACTAGTCCGGAGCTGGGC 296
QY 181 AAAGGCACCTATGGGAAGTTGACCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
DB 297 AAAGGCACCTATGGGAAGTTGACCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 356
QY 241 CTGAGGTTTGTGAACAAGCAAAACCAAGCTGAGAACTTCTTACGGGACGAGTGAGCATC 300
DB 357 CTGAGGTTTGTGAACAAGCAAAACCAAGCTGAGAACTTCTTACGGGACGAGTGAGCATC 416
QY 301 ACCAAGCAGCTCTCTCCAGCGCCCTTCATCATCAAGGCTTTTACGTGGTCTTTTCAGACA 360
DB 417 ACCAAGCAGCTCTCTCCAGCGCCCTTCATCATCAAGGCTTTTACGTGGTCTTTTCAGACA 476
QY 361 GAGGACTGCTACCTTTTGGCCAGGAGTACGACCTGTGGGACCTGTTTGCATCATC 420
DB 477 GAGGACTGCTACCTTTTGGCCAGGAGTACGACCTGTGGGACCTGTTTGCATCATC 536
QY 421 CCTCCCGAGGTGGGCTCCCTGAGACACAGGTGAGCGCTGTGTGACAGCTGGGCGCTG 480
DB 537 CCTCCCGAGGTGGGCTCCCTGAGGACACAGGTGAGCGCTGTGTGACAGCTGGGCGCTG 596
QY 481 GCGCTGGACTTCATGACGGGCGGAGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 540

DB 597 GCGCTGGACTTTCATGCAAGCGGCGGAGCTGTGTGACCGCGCATCAAGCGCGGAGAACGTG 656
QY 541 CTGCTGTTCGACCGCGAGTGCCTGCGGCTAAAGCTGGCCGACTTCGGCATCAGCGCGCGC 600
DB 657 CTGCTGTTCGACCGCGAGTGCCTGCGGCTAAAGCTGGCCGACTTCGGCATCAGCGCGCGC 716
QY 601 GTGGGCTGCCCGCTCAAGCGCGTGAAGCGGCGACCACTCCCTTACACGGGCGCTGAGGTGTC 660
DB 717 GTGGGCTGCCCGCTCAAGCGCGTGAAGCGGCGACCACTCCCTTACACGGGCGCTGAGGTGTC 776
QY 661 CAGCGGCGCGCGCGGAGTGGCTGGCGGTGACACAGCGGCGTGGAGCTGTGGGCTTGGC 720
DB 777 CAGCGGCGCGCGCGGAGTGGCTGGCGGTGACACAGCGGCGTGGAGCTGTGGGCTTGGC 836
QY 721 GTGCTCATCTTCTCGTGTCTCAGCGGCTACTTCCCGTGGGAGGCGGCTGGGCGCGCGAC 780
DB 837 GTGCTCATCTTCTCGTGTCTCAGCGGCTACTTCCCGTGGGAGGCGGCTGGGCGCGCGAC 896
QY 781 GCCTTCTTGAGAGTGTGCTGCTGCGTGGAGCGGCGGCGCTTCCCGGCGTCCCTTCCGAG 840
DB 897 GCCTTCTTGAGAGTGTGCTGCTGCGTGGAGCGGCGGCGCTTCCCGGCGTCCCTTCCGAG 956
QY 841 TGGCGCGCTTACCGGAGCGCGCTGCGCATGTTCAGCGCTTACTGGGCGCTGGAGCGC 900
DB 957 TGGCGCGCTTACCGGAGCGCGCTGCGCATGTTCAGCGCTTACTGGGCGCTGGAGCGC 1016
QY 901 GAGCGCGCGCGCGCGGAGGCTGTCGCTTCTCAAGCAGGAGCTCAGCTCCGAG 960
DB 1017 GAGCGCGCGCGCGGAGGCTGTCGCTTCTCAAGCAGGAGCTCAGCTCCGAG 1076
QY 961 CTGCGCGCGCGCGCTGCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGC 1020
DB 1077 CTGCGCGCGCGCGCTGCGACCGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGC 1136
QY 1021 GGGCCACTCGCTCGAGGCGCTGGGCGCGCTCAAGCGGAGCTGCTGACCGGAGCGGC 1080
DB 1137 GGGCCACTCGCTCGAGGCGCTGGGCGCGCTCAAGCGGAGCTGCTGACCGGAGCGGC 1196
QY 1081 AGCGGCTCCCGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG 1140
DB 1197 AGCGGCTCCCGCGCGCGCGCGCGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG 1256
QY 1141 GTGCGAGTGCCTGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG 1200
DB 1257 GTGCGAGTGCCTGCGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGG 1316
QY 1201 GGGCGGACCGAGCGCGCGCGGCAAGCAAGGAGGAGGAGTGTGTGGCGGCGCGCGCATC 1260
DB 1317 GGGCGGACCGAGCGCGCGGCAAGCAAGGAGGAGGAGTGTGTGGCGGCGCGCGCATC 1376
QY 1261 GAGATCTGCGTCTGA 1275
DB 1377 GAGATCTGCGTCTGA 1391
RESULT 2
AAS17862
ID AAS17862 standard; cDNA; 2598 BP.
XX AAS17862;
AC AAS17862;
XX 08-MAY-2002 (first entry)
XX Human protein kinase N cDNA.
DE Human; ss; protein kinase N; cytosolic; neuroprotective; cancer;
KW gene therapy; antigen; antibody; neurodegenerative disease;
KW inflammation; arteriosclerosis; psoriasis; growth disorder;
KW chromosome 16; papilloma virus infection, Alzheimer's disease.
OS Homo sapiens.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 14:17:18 ; Search time 299 Seconds
(without alignments)
11510.982 Million cell updates/sec

Title: US-10-016-985-1

Perfect score: 1275

Sequence: 1 atagagctgggtgccaga.....ccatcgagatctggtctga 1275

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1275	100.0	1473	24	ABK86975 Novel human protei
2	1275	100.0	2598	24	AA517862 Human protein kina
3	1273.4	99.9	1275	22	AA506744 Polynucleotide seq
4	1273.4	99.9	1275	25	AAAD49419 Human serine/threo
5	1273.4	99.9	2060	24	AAAD34298 Human PKIN-1 cDNA
6	1273.4	99.9	3124	25	AAAD49420 Human serine/threo
7	1262.4	99.0	1272	25	AAAD49416 Human serine/threo
8	1131.8	88.8	2893	24	ABK10101 Human cDNA encodin

9	963.4	75.6	1527	21	AAA09825 Rat PK5 gene, Rat
10	931.2	73.0	936	25	AA049417 Human serine/threo
11	849.6	66.6	7301	24	AA517863 Human protein kina
12	767.6	60.2	788	24	AA520990 DNA sequence #21 r
13	767.6	60.2	788	24	AA521038 Isolated CpG islan
14	736.6	18.6	1247	24	ABK89200 DNA encoding human
15	236.6	18.6	1716	22	AA506705 Polynucleotide seq
16	235	18.4	1156	24	AA034313 Human PKIN-16 cDNA
17	196	15.4	1119	22	AA506707 Polynucleotide seq
18	196	15.4	2380	25	AA049443 Human kinase and p
19	179	14.0	1887	25	AA049413 Human serine/threo
20	147.2	11.5	1865	23	AB117453 Drosophila melanog
21	147.2	11.5	14775	23	AB117452 Drosophila melanog
22	131.4	10.3	2132	23	ABU08603 Zea mays DNA fragm
23	105.8	8.3	1435	21	AA046591 Novel protein kina
24	100.4	7.9	1788	22	AAF44624 Human kinase (PKIN
25	99.8	7.8	1487	24	AA038855 Human kinase (PKIN
26	99.8	7.8	1640	22	AA086645 Human serine/threo
27	99.8	7.8	1864	24	ABQ76181 Human polynucleoti
28	99.8	7.8	2007	22	AA159743 Human NS cDNA sequ
29	99.6	7.8	2182	24	ABU39744 Human neuroblastom
30	98.4	7.7	774	22	AA197091 Human polynucleoti
31	98.2	7.7	1972	22	AA157962 Human Akt1 encodin
32	97.6	7.7	1443	24	ABU50837 PKB-green fluoresce
33	97.6	7.7	2181	19	AAV71037 Green fluorescent
34	97.6	7.7	2184	19	AAV71086 Human RAC protein
35	97.6	7.7	2610	18	AA064812 Human AKT-1 encodi
36	97.6	7.7	2610	18	AA067135 Human Akt-1 DNA se
37	97.6	7.7	2610	20	AAZ22190 Wild type human Ak
38	97.6	7.7	2610	21	AA09075 Human cDNA differe
39	97.6	7.7	2610	24	ABK84055 Human colon cancer
40	97.6	7.7	2610	24	AA028548 Human Akt2 encodin
41	97.6	7.6	718	21	AAH31058 Mouse Akt-2 cognat
42	96.8	7.3	1446	24	ABU50838 Human Akt-2 nucleo
43	93.4	7.3	1599	18	AA071252 Human Akt-2 nucleo
44	93.4	7.3	1599	21	AAA08443
45	93.4	7.3	1599	21	AAA08443

ALIGNMENTS

RESULT 1

ABK86975
ID ABK86975 standard; cDNA; 1473 BP.

AC ABK86975;

XX

XX 24-SEP-2002 (first entry)

DT Novel human protein (NHP) coding DNA.

DE Human; Novel human protein; NHP; gene; ss; kinase; regulatory pathway;

XX Therapeutic; diagnostic; pharmacogenomic; antigenic.

XX Homo sapiens.

XX Key

PH Location/Qualifiers

CDS 117..1391

FT /*tag= a

FT /product= "NHP"

FT /note= "This region is specifically claimed in claim 1 of the specification"

XX WO200246428-A2.

XX 13-JUN-2002.

PD 07-DEC-2001; 2001WO-US48533.

XX 07-DEC-2000; 2000US-251941P.

XX (LEXI-) LEXICON GENETICS INC.

Db 504 AGTACGTGCCCGGGGAGATGTTCTCACACCTACGGCGATCGGAAGTTCAAGTGC 563
Qy 446 ACACGGTGAAGCGTGTGTGACAGCTGGGCTTGGGCTGGACTTCATGACGGGGCG 505
Db 564 CCCATGCCCGTTTCTACGGCGCCAGATCGTCTGACCTTTGAGTATCTGCACTCGCTGG 623
Qy 506 AGCTGTGTACCGGACATCAAGCCCGAGACGTGCTGCTTTGACGGCGAGTGCAGCC 565
Db 624 ATCTCATCTACAGGACCTGAAGCCCGAGAACTGCTCAITGACAGCAGG-----GCT 677
Qy 566 GCGTAAAGCTGGCGGACTTCCGGCATGACGCCCGCTGGGCTGCGGCTCAAGCGCGTGA 625
Db 678 ACATTCAAGTACAGACTTCGTTTCCCAAGCGGCTGAAGGGCGGCACTTGGACCTTGT 737
Qy 526 GCGGACCATCCCTTACACGGCGCTGAGGTGTCCAGCGGGCGGCGCCGACGGGCTGG 685
Db 738 GCGGACCCCTGAGTACTGCGCCCTGAGATTATCTTGAGCAAAAGGCTACAAAGGCGG 797
Qy 686 CGGTGGACACGGCGTGGACGT 707
Db 798 TGGACTGTGGGCCCTGGGGT 819

RESULT 15

US-09-842-307-1
; Sequence 1, Application US/09842307
; Patent No. 6562574
; GENERAL INFORMATION:
; APPLICANT: Altschuler, David M.
; APPLICANT: Hirschhorn, Joel N.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: ASSOCIATION OF PROTEIN KINASE C ZETA
; TITLE OF INVENTION: POLYMORPHISMS WITH DIABETES
; FILE REFERENCE: 2825.2011-001
; CURRENT APPLICATION NUMBER: US/09/842.307
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,468
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2146
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-842-307-1

Query Match 5.8%; Score 73.8; DB 4; Length 2146;
Best Local Similarity 47.8%; Pred. No. 1.2e-05;
Matches 296; Conservative 0; Mismatches 302; Indels 21; Gaps 2;
Qy 152 AGCACTAGCACTAGTCCGGAGCTGGCAAGGACCACTATGGGAAGGTGACCTGGTGG 211
Db 731 AGGACTTTGACCTAATCAGAGTCATCGGGCGCGGAGCTACGCCAAGGTTCTCCTGGTGC 790
Qy 212 TCTACAAAGGGCACAGGCACAAATGACCTGAAGTTGTGAACA-----GAGCA 262
Db 791 GGTGAGAGAGAATGACCAATTTACGCCATGAAGTGTGAAGAAAGAGCTGGTGCATG 850
Qy 263 AAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATCACCAAGAGCTTCTCCAGCC 322
Db 851 ATGACGAGGATATTGACTGGGTACAGACAGAGAAGCAGCTGTTTGGCAGGCAATCCAGCA 910
Qy 323 CTTTCATCATCAAGGTCTTTGACGTGTTTGTGACAGAGGACTGTACGTCTTTGCC 382
Db 911 ACCCTTCTCGTGGTCCGATTAACACTCTGCTTCCAGACGACAAGTCGGTGTCTCTGTCA 970
Qy 383 AGGAGTACGACCTGCTGGGACCTGTTTGTGATCATCATCCCTCCAGGTGGGCTCCCTG 442
Db 971 TTGAGTACGTCAAGCGGGGACCTGATGTTTCCATGTGAGGAGCAGGAAGCTCCCTG 1030
Qy 443 AGGACAGGTGAAGCGCTGTGTGAGCAGAGCTGGGCTGGGCTGAGCTTTCATGACGGGC 502
Db 1031 AGGAGCAGCCAGGTTCCTACGCGCGGAGATCTGCATCGCCCTCAACTTCTCTGCACGAGA 1090

Qy 503 GGCAGCTGGTGCACCGGACATCAAGCCCGAGAACGTGCTGCTGTTTCGACCGCGAGTGCC 562
Db 1091 GGGGGATCATCTACAGGACCTGAAGCTGGAACACGTCTCTCTGATGCGGACGGGCACA 1150
Qy 563 GCGCGTAAAGCTGGCGGACTTCGGCATGACGGCGCGCTGGGCTGCGGCTCAAGCGCG 622
Db 1151 TCAAGCTCAAGACTACGGCATGTGCAAGGAAGGCTTGGGCCCCCTGGTGACACAACAGCA 1210
Qy 623 TGAGCGGCACCATCCCTTTACACGGCGCTGAGGTGTGCCAGGCGGCGCGCGCCACGGC 682
Db 1211 CTTTCTGCGGAACCCCGAATTACATCGCCCCCGAAATCTTGGGGAGAGAGTACGGGT 1270
Qy 683 TGGCGGTGGACACGGGCGGTGACGTGTGGGCTTTCGGCGTGTCTCATCTTCTGCGTGTCA 742
Db 1271 T-----CAGCGTGGACTGTGGGCGCTGGGAGTCTCTCATGTTTGAATGATGG 1318
Qy 743 CCGGCAACTTCCCGTGGGA 761
Db 1319 CCGGGCGCTCCCGTTTCA 1337

Search completed: December 7, 2003, 16:47:45
Job time : 84 secs

Db 445 AGTACGTGCCCGGGGAGATGTTCTCACACCTACGGCGGATCGGAAGGTTCAGTGAGC 504
QY 446 ACACGGTGAAGCGCTGTGTGAGCAGCTGGGCGCTGGGCTGGGACTTCATGACGGGGGGC 505
Db 505 CCATGCGCGGTTTCTACGCGGCCGAGATGCTCTGACCTTTGATATCTGCACCTCGTGG 564
QY 506 AGCTGTGACCGCGACATCAAGCCCGAGAACGTGTGCTGTTCGACCGGAGTGGCGCC 565
Db 565 ATCTCATCTACAGGACCTGAAGCGGAGATCTGCTCAITGACACGAGG-----GCT 618
QY 566 GCGTAAAGTGGCGGACTTCGGCATGACGCGCGCTGGGCTGGGCTCAAGCGGTGA 625
Db 619 ACATTGAGGTGACAGACTTCGGTTTCGCCAAGCGGTGAAGGGCCGCACTTGGACCTTGT 678
QY 626 GCGGCACCATCCCTTACACGCGCGCTGAGGTGTGCCAGCGCGCGCGCGCGCGGCTGG 685
Db 679 GCGGCACCTCTGACTCTGACCTGCGCTTTCGATATCTCTGAGCAAGGCTACACAGGCGG 738
QY 686 CGGTGGACACGGCGGTGACGT 707
Db 739 TGGACTGTGGCGCCCTGGGGT 760

RESULT 13

US-09-394-455-5
; Sequence 5, Application US/09394455
; Patent No. 6531305
; GENERAL INFORMATION:
; APPLICANT: Witman, George F.
; APPLICANT: San Agustín, Jovenal
; APPLICANT: Leszyk, John D.
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
; TITLE OF INVENTION: NUCLEIC ACIDS, AND METHODS OF USE
; FILE REFERENCE: 07917/078001
; CURRENT APPLICATION NUMBER: US/09/394,455
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/099,771
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)...(1133)
; NAME/KEY: misc feature
; LOCATION: (1)...(2549)
; OTHER INFORMATION: n = A,T,C or G

US-09-394-455-5

Query Match 5.8%; Score 74.4; DB 4; Length 2549;
Best Local Similarity 49.6%; Pred. No. 9,6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
QY 152 AGCTACGAACTAGTCCGGAGCTGGGCAAGGCAACCTATGGGAAGTTGACCTGTGG 211
Db 205 ATCAGTTTGAACGAATCAAGACCTCGGCACGGGCTCCTTCGGGCGGTGATGCTGTGA 264
QY 212 TCTAAGGGGACAGGACAAATGGCACTGAAGTTTGTGAACAGAGCAAAA---CCA 268
Db 265 AACACAGGAGACCGGGAACCACTATGCGCATGAAGATCTCTGACAAACAGAGGTGGTA 324
QY 269 AGCTGAAGAACTTCCTACGGGAGTGAGCATCAACAGACCTCTCCTCCAGCCCTTCA 328
Db 325 AACTGAACAGATCGAACACACCTCTGAATGAAGGCACTCTGCAAGCTGTCACTTTC 384
QY 329 TCATCAAGGTCTTTGACGTGTCTTTGACAGAGAGCTGTACGT---CTTTGCCAGG 385
Db 385 CGTTCTCGTCAAACTCGAGTTCTCCTTCAAGGACAACTCAAACTTATACATGTGTCTGG 444

QY 386 AGTACGCACTGTCTGGGAGCTGTTTGAATCATCTCCTCCAGGTGGGCTCCCTGAGG 445
Db 445 AGTACGTGCCCGGGGAGATGTTCTCACACCTACGGCGGATCGGAAGTTTCAGTGAGC 504
QY 446 ACACGGTGAAGCGCTGTGTGAGCAGCTGGGCGCTGGGCTGGGACTTCATGACGGGGGGC 505
Db 505 CCATGCGCGGTTTCTACGCGGCCGAGATGCTCTGACCTTTGATATCTGCACCTCGTGG 564
QY 506 AGCTGTGACCGCGACATCAAGCCCGAGAACGTGTGCTGTTCGACCGGAGTGGCGCC 565
Db 565 ATCTCATCTACAGGACCTGAAGCGGAGATCTGCTCAITGACACGAGG-----GCT 618
QY 566 GCGTAAAGTGGCGGACTTCGGCATGACGCGCGCTGGGCTGGGCTCAAGCGGTGA 625
Db 619 ACATTGAGGTGACAGACTTCGGTTTCGCCAAGCGGTGAAGGGCCGCACTTGGACCTTGT 678
QY 626 GCGGCACCATCCCTTACACGCGCGCTGAGGTGTGCCAGCGCGCGCGCGCGCGGCTGG 685
Db 679 GCGGCACCTCTGACTCTGACTCTGCGCTTTCGATATCTCTGAGCAAGGCTACACAGGCGG 738
QY 686 CGGTGGACACGGCGGTGACGT 707
Db 739 TGGACTGTGGCGCCCTGGGGT 760

RESULT 14

US-09-394-455-35
; Sequence 35, Application US/09394455
; Patent No. 6531305
; GENERAL INFORMATION:
; APPLICANT: Witman, George F.
; APPLICANT: San Agustín, Jovenal
; APPLICANT: Leszyk, John D.
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
; TITLE OF INVENTION: NUCLEIC ACIDS, AND METHODS OF USE
; FILE REFERENCE: 07917/078001
; CURRENT APPLICATION NUMBER: US/09/394,455
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 60/099,771
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 2608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)...(1192)

US-09-394-455-35

Query Match 5.8%; Score 74.4; DB 4; Length 2608;
Best Local Similarity 49.6%; Pred. No. 9,6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
QY 152 AGCTACGAACTAGTCCGGAGCTGGGCAAGGCAACCTATGGGAAGTTGACCTGTGG 211
Db 264 ATCAGTTTGAACGAATCAAGACCTCGGCACGGGCTCCTTCGGGCGGTGATGCTGTGA 323
QY 212 TCTAAGGGGACAGGACAAATGGCACTGAAGTTTGTGAACAGAGCAAAA---CCA 268
Db 324 AACACAGGAGACCGGGAACCACTATGCGCATGAAGATCTCTGACAAACAGAGGTGGTA 383
QY 269 AGCTGAAGAACTTCCTACGGGAGTGAGCATCAACAGACCTCTCCTCCAGCCCTTCA 328
Db 384 AACTGAACAGATCGAACACACCTCTGAATGAAGGCACTCTGCAAGCTGTCACTTTC 443
QY 329 TCATCAAGGTCTTTGACGTGTCTTTGACAGAGAGCTGTACGT---CTTTGCCAGG 385
Db 444 CGTTCTCGTCAAACTCGAGTTCTCCTTCAAGGACAACTCAAACTTATACATGTGTATGG 503
QY 386 AGTACGCACTGTCTGGGAGCTGTTTGAATCATCTCCTCCAGGTGGGCTCCCTGAGG 445

Db 124 AGRAGAATGATACCAAGAGATGCGCAATGAAGTATACATGAATAAACAAGTGGCTGG 183
Qy 272 TGAAGAACTTCCCTACGGAGGTGAGCA-----TCACCAACAGCCTCTCCCTCCAGCCCT 325
Db 184 AGCGCAATGAAGTGAAGAAATGCTTCAAGGAATCCAGATCATGACGGGTCTGGAGCACC 243
Qy 326 TCATCATCAAGGTCTTTGACGTGCTTTGAGACAGAGGACTGCTAGTCTTTCCCCAGG 385
Db 244 CTTTCTCGTTAAATTTGGTATTCCTTCCAAAGTAGGAAGACATGTTCAATGTTGGTGG 303
Qy 386 AGTACGACACCTGCTGGGACCTGTTGACATCATCCCTCCCGAGTGGGCTCCCTGAGG 445
Db 304 ACCTCTCTGGTGGAGACCTGCTTATCACCTGCAACAGAGCTCCACTTCAAGNAG 363
Qy 446 ACAGGTGAAGCGCTGTGTGAGCAGCTGGGCTGGGCTGGACTTTATGACCGGGGGC 505
Db 364 AAACAGTGAAGCTTCTATCTGTGAGCTGTGATGGGCTGGACTACTGTCAGAACACG 423
Qy 506 AGCTGGTGCACCGGACATCAAGCCCGAGACGCTGCTCTTGCAGCGGAGTGGCCGCC 565
Db 424 GCATCATTCACAGGATATGAAGCTGACATATTTTACTTGAAGAACATGGGACGTGC 483
Qy 566 GCGTAAAGCTGGCGACTTCGGCATGACGCGCGGCTGGGCTGGCGGTCAAGCGCGTGA 625
Db 484 ACATCAGAGATTTCAACATTCTCGATGCTGCCAGGAGACACAGATTACCACTGG 543
Qy 626 GCGGACACATCCCTACAGCGGCTGAGGTGTC 660
Db 544 CTGGCACCAAGCCTTACATGGCACTGAGATGTTTC 578

RESULT 11
US-09-394-455-3
; Sequence 3, Application US/09394455
; Patent No. 6531305
; GENERAL INFORMATION:
; APPLICANT: Witman, George F.
; APPLICANT: San Agustin, Jovenal
; APPLICANT: Leszyk, John D.
; TITLE OF INVENTION: SPERM ASSOCIATED PROTEIN KINASE POLYPEPTIDES, CORRESPONDING
; FILE REFERENCE: 07917/078001
; CURRENT APPLICATION NUMBER: US/09/394,455
; PRIOR FILING DATE: 1999-09-10
; PRIOR FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1008
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1008)
US-09-394-455-3

Query Match 5.8%; Score 74.4; DB 4; Length 1008;
Best Local Similarity 49.6%; Pred. No. 8.6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
Qy 152 AGCACTAGCACTAGTCCGGAGTGGGCAAGGACCTATGGAGGTTGACCTGGTGG 211
Db 80 ATCAGTTTGAACGAATCAAGACCTTCGGCAGCGGCTCCTTCGGCGGGTGTGTTGGTGA 139
Qy 212 TCTAAGGACACAGGACACAAAATGGCACTGAAGTTTGTGAACAAGACAAA---CCA 268
Db 140 AACCAAGGACCGGGAACCACTATGCGATGAAGATCCTTCGACNAACAGAGGTTGTA 199
Qy 269 AGCTGAAGAACTTCTTACGGAGGTGAGCATCACCAAGCCTCTCTCCAGCCCTTCA 328
Db 200 AACTGAACACAGATGAAACACACCCCTGATGAAGAACGCGATCTCTGCAAGCTGCAACTTC 259

Qy 329 TCATCAAGGTCTTTGAGCTGCTTTTGAAGACAGAGGACTGCTAGCT---CTTTGCCAGG 385
Db 260 CGTTCCTCGTCAAACTCGAGTTCTCCTTCAAGGACAACTCAAACTTATACATGTCATGG 319
Qy 386 AGTACGACACCTGCTGGGACCTGTTTGAATCATCCCTCCCGAGTGGGCTCCCTGAGG 445
Db 320 AGTACGTGCGCGGGGAGATGTTCTCACACTACGCGGATCGGAAGGTTCACTGAGC 379
Qy 446 ACAGGTGAAGCGCTGTGTGACAGCTGGGCTTGGGCTGGACTTCATGACAGGCGGC 505
Db 380 CCATGCGCGCTTCTACGCGCCAGATCGTCTCTGACCTTTGATGATCTGCACTCGCTGG 439
Qy 506 AGCTGTGACCGGACATCAAGCCCGAGAGCTGTGCTGTGTCGACCGGAGTGGCGCC 565
Db 440 ATCTCATCTACAGGACCTGAAGCGGAGAACTCTGCTCATTTGACCACGAG---GCT 493
Qy 566 GCGTAAAGCTGGCGACTTCGGCATGACGCGCGGCTGGGCTGGCGGTCAAGCGCGTGA 625
Db 494 ACATTGAGTGAACAGCTTCGTTTCGCAAGCGGTGAAGGCGGCACTTGGACCTTGT 553
Qy 626 GCGGACACATCCCTTACAGCGGCGCTGAGTGTGCCAGGCGGCGCGCGACAGGCTGG 685
Db 554 GCGGACACCTGAGTACCTGCGCCCTGAGATTATCTCTGAGCAAGGCTCAACAAGGCCG 613
Qy 686 CCGTGGACACGGGCGTGGACGT 707
Db 614 TGGACTGGTGGGCGCTGGGGT 635

RESULT 12
US-09-467-082-3
; Sequence 3, Application US/09467082
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF PKA CATALYTIC SUBUNIT C-ALPHA EXPRESSION
; FILE REFERENCE: RTS-0088
; CURRENT APPLICATION NUMBER: US/09/467,082
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 3
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(1136)
; NAME/KEY: unsure
; LOCATION: 6
; OTHER INFORMATION: unknown
US-09-467-082-3

Query Match 5.8%; Score 74.4; DB 3; Length 2549;
Best Local Similarity 49.6%; Pred. No. 9.6e-06;
Matches 279; Conservative 0; Mismatches 271; Indels 12; Gaps 3;
Qy 152 AGCACTAGCACTAGTCCGGAGTGGGCAAGGACCTATGGAGGTTGACCTGGTGG 211
Db 205 ATCAGTTTGAACGAATCAAGACCTTCGGCAGCGGCTCCTTCGGCGGGTGTGTTGGTGA 264
Qy 212 TCTAAGGACACAGGACACAAAATGGCACTGAAGTTTGTGAACAAGACAAA---CCA 268
Db 265 AACCAAGGACCGGGAACCACTATGCGATGAAGATCCTTCGACAAACAGAGGTTGTA 324
Qy 269 AGCTGAAGAACTTCTTACGGGAGTGAAGATCACCACAGCCTCTCTCCAGCCCTTCA 328
Db 325 AACTGAACAGATGCAACACACCTGATGAAGGCGCATCTCTGCAAGCTGTCACCTTC 384
Qy 329 TCATCAAGGTCTTTGAGCTGCTTTTGAAGACAGGACTGCTAGCT---CTTTGCCAGG 385
Db 385 CGTTCCTCGTCAAACTCGAGTTCTCCTTCAAGGACAACTCAAACTTATACATGTCATGG 444
Qy 386 AGTACGACACCTGCTGGGACCTGTTTGAATCATCCCTCCCGAGGTTGGGCTCCCTGAGG 445

Db 540 CTTGCACTATCTCAAACTCCTTGGCAAGGAAACCTTTGGCAAAAGTCATCTGTGGCGGA 599
Qy 216 CAAGGGCACAGGACAAATAATGGCACTGAAGTTTGTGAAACAAGAGCAAAACCAAGCTGAA 275
Db 600 GAAGGCCACTGGCCGCTACTAGCCATGAGATCCTGGAAAGGAAGTCTCATTTGCCAA 659
Qy 276 GAATCTCTTACGGAGGTGAGTACACCAACAG-----CCTCTCTCCAGCCCTTCAT 329
Db 660 GGATGAAGTGGCTCACACAGTCAACGAGAGCGGGTCTCTCAGAAACACCAAGGACCCCGTT 719
Qy 330 CATCAAGGTCTTTGACGTGCTTTGACAGAGAGTCTTACGTCTTTGCCAGAGTA 389
Db 720 CCTCACTGCGTGAAGTATGCTTCCAGACCCAGCCGCTGTGCTTTGATGAGTA 779
Qy 390 CGCACTGCTGGGACCTGTTTGACATCATCCCTCCACAGGTGGGGTCTCCTGAGGACAC 449
Db 780 TGCCAAACGGGGTGAGCTGTTCTTCCACCTGTCGGGAGCGTCTTTCACAGAGAGCG 839
Qy 450 GGTGAAGCGCTGTGTGACAGAGCTGGGCTTGGCTGTGACTTGTATGACAGGGGCGGAGCT 509
Db 840 GGCCCGGTTTATGGTGACAGATTTGCTCGGCTCTTGAAGTCTTGCATCTGCACTCGCGGACGT 899
Qy 510 GGTGACCGGCACATCAAGCCGAGAGAGTCTGCTGTTGACCGCGAGTGCCTCGCGT 569
Db 900 GGTATACCGGACATCAAGCTGGAACACCTCATGCTGACAAAGATGGCCACATCAAGAT 959
Qy 570 AAAGCTGCGGACATCGGCAATGACCGCGCGGTGGGCTGCGGCTCAAGCGCGTGAGCGG 629
Db 960 CACTGACTTTGGGCTCTGCAAGAGGACATCAGTGACGGGCGCCACCATGAAACCTTCTG 1019
Qy 630 CACCATCCTTACAGGCGCTGAGGTGTGCGAGCGGCGCGCGCGGCGGCGGCTGGCGGT 689
Db 1020 TGGACCGCGGAGTACCTTGGCGGCTGAGGTGCTGAGGACATGACTATGGCGCGGCGGT 1079
Qy 690 GGACAGCGGCTGAGCGTGTGGGCTTGGGCTGCTGCTATCTTGTGCTGCTCACCGCAA 749
Db 1080 GGACTGTGGGGGCTGGGTGGTGTGATGACAGATGATGCGCGCGCGCTTGGCTTCTA 1139
Qy 750 CTTCCGCTGGAGGCGGCGTGGGCGCGGCGGCGGCGGCTTCTTGGAGAGTTTC 798
Db 1140 CAACGAGGACACGAGCGGCTTTCGAGCTCATCTCATGGAAGATC 1188

RESULT 9

US-09-428-711A-15
; Sequence 15, Application US/09428711A
; Patent No. 6358720
; GENERAL INFORMATION:
; APPLICANT: Muramatsu, Masaaki
; APPLICANT: Shirasawa, Takuji
; APPLICANT: Tokumitsu, Hiroshi
; APPLICANT: No. 6358720uchi, Teruhisa
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE
; FILE REFERENCE: 06501-045001
; CURRENT APPLICATION NUMBER: US/09/428,711A
; CURRENT FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: PCT/JP98/01246
; PRIOR FILING DATE: 1998-03-23
; PRIOR APPLICATION NUMBER: JP 9/124798
; PRIOR FILING DATE: 1997-04-28
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 5228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (269)... (3418)
US-09-428-711A-15

Query Match 6.4%; Score 82.2; DB 4; Length 5228;
Best Local Similarity 46.4%; Pred. No. 3.6e-07;

Matches 267; Conservative 0; Mismatches 308; Indels 0; Gaps 0;
Qy 236 TGGCACTGAAGTTTGTGAAACAAGCAAGCACTGAAGAACTTCTTACGGAGGTGA 295
Db 396 TCGCCGTCAGTGCATTAACAAGAGAACTCTGCCAAGTCTCAGACGCTGCTGGGAAGG 455
Qy 296 GATCAACCAACAGCCTCTCTCCAGCCCTTCAATCATCAAGGTCTTTGAGTGGTCTTTG 355
Db 456 AAATCAAAATCTTGAAGAACTGAAACAATGAAGAACTCTGTGGCCCTGTAGACTTCCAGG 515
Qy 356 AGACAGAGGACTCTAGCTTTTGGCCCAAGAGTACGACCTGCTGGGAGCTCTTTTGACA 415
Db 516 AATAGGCTAATTTCTGTCTTACCTGTTATGAGTACTGCAACGGTGGGAGCTTGGCCGACT 575
Qy 416 TCATCCCTCTCCCAAGTGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGTCAGCAGCTGG 475
Db 576 ACCTGCAGGCCATGCGCAGCTGAGCGAGGACACCATCAGGCTCTTCTCTCAGCAGATCG 635
Qy 476 GCTTGGCGCTGGAGCTTATGACAGGGGGGAGCTGTGTGCAACCGGACATCAAGCCCGAGA 535
Db 636 CGGCGCCATGCGGCTTCTGCACAGCAAGGCAATCATCCACCGGACCTGAAACCCGAGA 695
Qy 536 ACCTGCTGCTGTTCGACCGGAGTGGCGCGCTAAAGCTGGCGACTTCGGCATGACGC 595
Db 696 ACATCTCTGCTTCAACCCCGCGCGCGCGCCCAACCCCAACAGCATCCGCGTCAAGA 755
Qy 596 GCGCGCTGGGCTGCGCGCTCAAGCGGTGAGCGGCAACCATCCCTTACA CGGCGCTGAGG 655
Db 756 TCGTGAATTCGGCTTCGCGCGTACTCTCCAGAGCAACATGATGGCGGCCACACTCTGCG 815
Qy 656 TGTGCCAGGCGCGCGCGCGAGCGGCTGGCGGTGACACGGCGGTGGAGCTGTGGGCT 715
Db 816 GCTCCCCCATGTATGCGCCCCCGAGTCTCATGTGCCAGCACTACGACGGGAAGCGG 875
Qy 716 TCGCGGTGCTCATCTTCTGCGTCTCACCGCAACTTCCCGTGGAGGCGCGCTCGGGCG 775
Db 876 ACTGTGAGCATCGGACCATGCTGTACAGTGCCTGACGGGAGGCGGCCCTTCCAGG 935
Qy 776 CCGAGCGCTTCTTTCGAGAGTTCGTGCGTGGCAG 810
Db 936 CCAGCAGCCCCAGGACCTCGCGCTGTCTACAG 970

RESULT 10

US-09-801-876B-1
; Sequence 1, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001160
; CURRENT APPLICATION NUMBER: US/09/801,876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Human.
US-09-801-876B-1

Query Match 5.9%; Score 75.4; DB 4; Length 1485;
Best Local Similarity 48.2%; Pred. No. 5.8e-06;
Matches 248; Conservative 0; Mismatches 261; Indels 6; Gaps 1;

Qy 152 AGCACTAGCACTAGTCCGGAGCTGGCAAGGACCTATGGGAAGGTTGACCTGTGG 211
Db 64 ACCACTTTGAAATTTTCGAGCCATTCGGAAGGCACTTTGGGGAGGCTCTGCAATTGAC 123
Qy 212 TCTCAAGGCGACAGGCACAAAATGGCACTGAAGTTTGTGAACAGAGCAAAACCAAGC 271

QY 609 CCGGTCAAGCGGTGAGCGCACCATCCCTTACACGGCGCTGAGGTGTGCGAGCGG 668
Db 1107 TGCACCATGAAGACCTTTTTCGGGCACACCTGAGTACCTTGCCCGGAGGTGCTGGAGGA 1166
QY 669 CCGCGCCGACGGGTGGCGGTGGACACGGGGGTGGAGTGTGGGCCCTTGGCGGTGCTCAT 728
Db 1167 CAATGACTACGGCGGTGAGTGGAGTGTGGGGGTGGCGGTGCTCATGTACAGATGAT 1226
QY 729 CTTCTGCGTGTCTACCGCGCAACTTCCCGTGGGAGCGCGCTCGGGCGCGGACGCGCTTCTT 788
Db 1227 GTGCGTGCCTGCCCTTCTACACAGGACCATGAGAAGCTTTTGTAGCTCATCTCAT 1286
QY 789 CGAGGAGTTC 798
Db 1287 GGAGGAGATC 1296

RESULT 7
US-09-256-465-1
; Sequence 1, Application US/09256465
; Patent No. 6043090
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-2 EXPRESSION
; FILE REFERENCE: RTS-0035
; CURRENT APPLICATION NUMBER: US/09/256.465
; CURRENT FILING DATE: 1999-02-23
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 1599
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(1533)
US-09-256-465-1

Query Match
Best Local Similarity 7.3%; Score 93.4; DB 3; Length 1599;
Matches 332; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

QY 96 CACTGAAGACATGAGCGCCCTGACTCTCCGACACTGCGCCCGGAGCGTACCAAGCA 155
Db 480 CACGACTGAGGAGTGGGAAGTGGCGGTGAGCAAGCGACGCGCTAAAGTGACCATGAATGA 539
QY 156 CTACGAACTAGTCCGGGAGCTGGGCAAGGACCTATGGGAAGTGTGACCTGGTGGTCTA 215
Db 540 CTTGCACTATCTCAAACTCTTGGCAAGGGAACCTTTGGCAAGTCTATCTTGGTGGCGGA 599
QY 216 CAAGGCACAGGCACAAATAATGGCACTGAAGTTGTGAACAGAGCAAAACCAAGCTGAA 275
Db 600 GAAGGCCACTGGCGCTACTACGCCATGAAGATCTGCGAAGGAAGTCAATTCATTTGCCAA 659
QY 276 GAATCTCTACGGAGGTGACATCAACCAAG-----CCTCTCTCCAGCCCTTCAT 329
Db 660 GGATGAAGTCTCTACAGTCAACGAGCCGGGTCTCCAGAACACACCGGACCCCGTT 719
QY 330 CATCAAGGTCTTTGACGTGGTCTTTGACAGAGGACTGTACGTCTTTTGCCAGGAGTA 389
Db 720 CTTCACTGCGTGAAGTATGCTTCCAGACCCAGCCGCTGTGCTTTGTGTGAGTA 779
QY 390 GCGACCTCTGGGACCTGTTTGACATCATCCCTCCCGAGGTGGGGTCTCCGTGAGGACAC 449
Db 780 TGCCAAACGGGGTGAAGTGTCTTCCAGACCCAGCCGCTGTGCTTTGTGTGAGTA 779
QY 450 GGTGAAGCGCTGTGACAGAGCTGGGCTGGCGTGGACTTCATGACAGGCGGCGAGCT 509
Db 840 GGCCCGGTTTATGGTCAGAGATGTCTCGGCTCTTGAGTACTTGCACTCGCGGAGCT 899
QY 510 GGTGACCGGACATCAAGCCGAGAGCGTGTGCTGTTCGACCGGAGTGCCTCCCGCT 569
Db 900 GGTATACCGGACATCAAGCTGGAAACCTCATGCTTGACAAAGATGGCCACATCAAGT 959

RESULT 8

US-09-167-322-3
; Sequence 3, Application US/09167322
; Patent No. 6365151
; GENERAL INFORMATION:
; APPLICANT: Allegheny University of the Health
; Sciences, Halpern, Michael S.
; England, James M.
; TITLE OF INVENTION: CANCER VACCINE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Suite 1800, Two Penn Center Plaza
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,322
; FILING DATE: 07-Oct-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/00582
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 7933-33 PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1599 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-167-322-3

Query Match 7.3%; Score 93.4; DB 4; Length 1599;
Best Local Similarity 46.8%; Pred. No. 2.5e-09;
Matches 332; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

QY 96 CACTGAAGACATGAGCGCCCTGACTCTCCGACACTGCGCCCGGAGCGTACCAAGCA 155
Db 480 CACGACTGAGGAGTGGGAAGTGGCGGTGAGCAAGCGACGCGCTAAAGTGACCATGAATGA 539
QY 156 CTACGAACTAGTCCGGGAGCTGGGCAAGGACCTATGGGAAGTGTGACCTGGTGGTCTA 215

QY 729 CTTCTGCTGCTCACCGGCAACTTCCGTCGGAGCGCGCTCGGGCCGCCGACGCTTCTT 788
Db 1770 GTGGTGCCTGCTCCCTTCTACACCAAGGACCATGAGAAGCTTTTGTAGCTCATCTCTCAT 1829
QY 789 CGAGGATTC 798
Db 1830 GGAGGAGATC 1839

RESULT 5

US-09-212-771-1
; Sequence 1, Application US/09212771
; Patent No. 5958773
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowert
; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-1 EXPRESSION
; FILE REFERENCE: RTS-0034
; CURRENT APPLICATION NUMBER: US/09/212,771
; CURRENT FILING DATE: 1998-12-16
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(1641)
US-09-212-771-1

Query Match 7.7%; Score 97.6; DB 2; Length 2610;

Best Local Similarity 48.8%; Pred. No. 4.3e-10;

Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;

QY 138 CAGCGAGCTCACCGAGCTACGAACTAGTCCGGAGCTGGGCAAGGACCTATGGAA 197
Db 627 CCGGTGACCATGAACGAGTTTGAGTACCTGCTGGCAAGGCACTTTCCGCAA 686
QY 198 GGTTCAGCTGCTGTCTACAAAGGCGACAGGCAAAAATGGCACTGAAGTTTGTGAACA 257
Db 587 GGTGATCTGTGAGGAGAGGCGACAGGCGCTACTACGCCATGAATCCTCAAGAA 746
QY 258 GAGCAAAACCAAGCTGAAGAACTTCCCTACGGGAGGTGAGCATCAACACGCTCTC --- 314
Db 747 GGAAGTCATCTGTGCAAGGACGAGGTGGCCACACACTCACCAGAAACCGCTCTGCA 806
QY 315 ---CTCCAGCCCTTCATCACTCAAGGCTTTGACGTGCTTTGAGACAGAGGACTGCTA 371
Db 807 GAACTCCAGGCAACCCCTTCTCAAGCCCTGAAGTACTCTTTCCAGACCCACGACCGCT 866
QY 372 CGTCTTTGCCAGGAGTACGACCTGCTGGGACCTGTTTGACATCATCCCTCCCGAGGT 431
Db 867 CTGCTTTGTATGGAGTACGCCACGCGGGGCGAGCTGTTCTTCCACTGTCCCGGAACG 926
QY 432 GGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACGAGCTGGGCGCTGGAATT 491
Db 927 TGTGTTCTCCAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTGAGCCCTGGACTA 986
QY 492 CATGCAC ---GGGCGGAGCTGTGACACCGGACATCAAGCCCGAGAACCTGTGCTGTT 548
Db 987 CTTGCACTCGGAGAAAGACGTGGTGTACCGGACCTCAAGTGGAGAACCTCATGCTGGA 1046
QY 549 CGACCGGAGTCCCGCCGCTAAAGCTGGCGGACCTTCGGCATGACCGCCCGCTGGGCTG 608
Db 1047 CAAGGACGGGACATTAAGATCACAGACTTCGGGCTGTGCAAGGAGGGATCAAGACGG 1106
QY 609 CCGGCTCAAGCGCGTGTGAGGCAACATCCCTTACACGGCGCTGAGGTGTGCCAGCGGG 668
Db 1107 TGCCACCATGAAGACCTTTTGGGACACACTGAGTACCTGGCCCGGAGGTGCTGGAGGA 1166
QY 669 CCGGCGGAGCGGCTGGCGGTGGACACGCGGCTGACAGTGTGGGCTTCGGGCTGCTCAT 728

Db 1167 CAATGACTACGGCCGTGTCAGTGGACTGTTGGGGCTTGGCGTGTCTATGTACGAGATGAT 1226
QY 729 CTTCTGCTGCTCACCGGCAACTTCCGTCGGAGCGCGCTCGGGCCGCCGACGCTTCTT 788
Db 1227 GTCCGTCGCTGCCCTTCTACAAACGAGGACCATGAGAAGCTTTTGTAGCTCATCTCAT 1286
QY 789 CGAGGATTC 798
Db 1287 GGAGGAGATC 1296

RESULT 6

US-09-091-058-1
; Sequence 1, Application US/09091058
; Patent No. 6054285
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Frech, Matthias
; TITLE OF INVENTION: Screening Method
; FILE REFERENCE: 4-20683/A/20684/PCT
; CURRENT APPLICATION NUMBER: US/09/091,058
; CURRENT FILING DATE: 1998-06-10
; EARLIER APPLICATION NUMBER: PCT/EP96/04814
; EARLIER FILING DATE: 1996-11-05
; EARLIER APPLICATION NUMBER: 9525703.6
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(1641)
US-09-091-058-1

Query Match 7.7%; Score 97.6; DB 3; Length 2610;

Best Local Similarity 48.8%; Pred. No. 4.3e-10;

Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;

QY 138 CAGCGAGCTCACCGAGCTACGAACTAGTCCGGAGCTGGGCAAGGACCTATGGAA 197
Db 627 CCGGTGACCATGAACGAGTTTGAGTACCTGCTGGCAAGGCACTTTCCGCAA 686
QY 198 GGTTCAGCTGCTGTCTACAAAGGCGACAGGCAAAAATGGCACTGAAGTTTGTGAACA 257
Db 587 GGTGATCTGTGAGGAGAGGCGACAGGCGCTACTACGCCATGAATCCTCAAGAA 746
QY 258 GAGCAAAACCAAGCTGAAGAACTTCCCTACGGGAGGTGAGCATCAACACGCTCTC --- 314
Db 747 GGAAGTCATCTGTGCAAGGACGAGGTGGCCACACACTCACCAGAAACCGCTCTGCA 806
QY 315 ---CTCCAGCCCTTCTATCACTCAAGGCTTTGACGTGCTTTGAGACAGAGGACTGCTA 371
Db 807 GAACTCCAGGCAACCCCTTCTCAAGCCCTGAAGTACTCTTTCCAGACCCACGACCGCT 866
QY 372 CGTCTTTGCCAGGAGTACGACCTGCTGGGACCTGTTTGACATCATCCCTCCCGAGGT 431
Db 867 CTGCTTTGTATGGAGTACGCCACGCGGGGCGAGCTGTTCTTCCACTGTCCCGGAACG 926
QY 432 GGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACGAGCTGGGCGCTGGAATT 491
Db 927 TGTGTTCTCCAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTGAGCCCTGGACTA 986
QY 492 CATGCAC ---GGGCGGAGCTGTGACACCGGACATCAAGCCCGAGAACCTGTGCTGTT 548
Db 987 CTTGCACTCGGAGAAAGACGTGGTGTACCGGACCTCAAGTGGAGAACCTCATGCTGGA 1046
QY 549 CGACCGGAGTCCCGCCGCTAAAGCTGGCGGACCTTCGGCATGACCGCCCGCTGGGCTG 608
Db 1047 CAAGGACGGGACATTAAGATCACAGACTTCGGGCTGTGCAAGGAGGGATCAAGACGG 1106

```
RESULT 3
US-09-417-197-70
; Sequence 70, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 70
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PKB-EGFP fusion
; NAME/KEY: CDS
; LOCATION: (1)..(2178)
US-09-417-197-70

Query Match      7.7%; Score 97.6; DB 4; Length 2181;
Best Local Similarity 48.8%; Pred. No. 4.2e-10;
Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;

Qy 138 CAGCGAGTCCACCAAGCACTACGAACTAGTCGGGAGCTGGGCAAGGCACTATGGAA 197
Db 429 CCGGCTGACCATGAACGAGTTTGAGTACCTGAAGCTGCTGGGCAAGGCACTTTCCGCAA 488

Qy 198 GGTTCACCTGTGTCTACAAAGGCGACAGGACAAAATGCACTGAAGTTTGTGAACAA 257
Db 489 GGTGATCTGTGTGAAGGAGGAGGCAAGGCGCTACTACGCCATGAAGATCTCTCAAGAA 548

Qy 258 GAGCAAAACCAAGTGAAGAACTTCTCAAGGAGGTGAGCATACCAACAGCCTCTC--- 314
Db 549 GGAAGTCACTGTGGCCAAAGGACGAGGTGGCCCAACACTACCGAGAACCGCGTCTGCA 608

Qy 315 ---CTCCAGCCCTTCATCATCAAGTCTTTGAGTGTCTTTGAGACAGAGACTGCTA 371
Db 609 GAATCCAGGACCCCTTCTCCACAGCCCTGAAGTACTCTTTCCAGACCCACGACCGCCT 668

Qy 372 CGTCTTTGCCAGGAGTACGCACTGTCTGGGACCTGTGTGACATCATCCTCCCGAGGT 491
Db 729 TGTGTTCTCCAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTGAGCCCTGGACTA 788

Qy 492 CATGCAC---GGCGGAGCTGTGTGACCGGACATCAAGCCCGAGGAGTGTGCTGTT 548
Db 789 CCTGCATCGGAGAAACGTTGTACCGGACCTCAAGGTGAGAACTCATGCTGGA 848

Qy 549 CGACCGGAGTGGCCCGCGCTAAAGCTGGCCGACTTCCGATGACGCGCCGCTGGGCTG 608
Db 849 CAAGGACGGGCACATTAGATCAAGACTTGGGCTGTGCAAGGAGGGGATCAAGGACGG 908

Qy 609 CCGGTCAAGCGGTGAGCGGACCATCCCTTTACACGGCGCTGAGTGTGCCAGGCGGG 668
Db 909 TGCACCATGAAGACCTTTTGGCGCACACCTGAGTACCTGCGCCCGAGGTGCTGGAGGA 968

Qy 669 CCGCGCCAGCGGTGGCGGTGGAACACGGCGTGAACGTGTGGCGCTTCGCGGTGCTCAT 728
Db 969 CAATGACTACCGCGCTGAGTGAACGTGTGGGGCTGGGCGCTGATGTACGAGATGAT 1028

Qy 729 CTTCTGCGTGTCAACGGCAACTTCCCGTGGGAGCGCGCTCGGCGCGCGCTTCTT 788
Db 1029 GTGGGTGCGCTGCGCTTCTCAACACGAGGACCATGAGAAGCTTTTGTGAGCTCATCTCAT 1088
```

```
Qy 789 CGAGGAGTTC 798
Db 1089 GGAGGAGATC 1098

RESULT 4
US-09-417-197-138
; Sequence 138, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An I
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 138
; LENGTH: 2184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-PKB fusion
; NAME/KEY: CDS
; LOCATION: (1)..(2181)
US-09-417-197-138

Query Match      7.7%; Score 97.6; DB 4; Length 2184;
Best Local Similarity 48.8%; Pred. No. 4.2e-10;
Matches 327; Conservative 0; Mismatches 334; Indels 9; Gaps 2;

Qy 138 CAGCGAGTCCACCAAGCACTACGAACTAGTCGGGAGCTGGGCAAGGCACTATGGAA 197
Db 1170 CCGGCTGACCATGAACGAGTTTGAGTACCTGAAGCTGCTGGGCAAGGCACTTTCCGCAA 1229

Qy 198 GGTTCACCTGTGTCTACAAAGGCGACAGGCAAAAATGCACTGAAGTTTGTGAACAA 257
Db 1230 GGTGATCTGTGTGAAGGAGGAGGCAAGGCGCTACTACGCCATGAAGATCTCTCAAGAA 1289

Qy 258 GAGCAAAACCAAGTGAAGAACTTCTCAAGGAGGTGAGCATACCAACAGCCTCTC--- 314
Db 1290 GGAAGTCACTGTGGCCAAAGGACGAGGTGGCCCAACACTACCGAGAACCGCGTCTGCA 1349

Qy 315 ---CTCCAGCCCTTCATCATCAAGTCTTTGAGTGTCTTTGAGACAGAGACTGCTA 371
Db 1350 GAATCCAGGACCCCTTCTCCACAGCCCTGAAGTACTCTTTCCAGACCCACGACCGCCT 1409

Qy 372 CGTCTTTGCCAGGAGTACGCACTGTCTGGGACCTGTGTGACATCATCCTCCCGAGGT 431
Db 1410 CTGCTTTGTGATGGAGTACGCCAACGCGGGCGAGCTGTCTTCCACTGTCCCGGGAACG 1469

Qy 432 GGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCAGCTGGGCGCTGGGCTGGAATT 491
Db 1470 TGTGTTCTCCGAGGACCGGGCCCGCTTCTATGGCGCTGAGATTGTGTGAGCCCTGGACTA 1529

Qy 492 CATGCAC---GGCGGAGCTGTGTGACCGGACATCAAGCCCGAGGAACTGTGCTGTT 548
Db 1530 CTTGCATCTCGAGAGAAACGTGTGTACCGGACCTCAAGCTGAGAACTCATGCTGGA 1589

Qy 549 CGACCGGAGTGGCGCGCTAAAGCTGGCCGACTTGGCATGACGCGCGCTGGGCTG 608
Db 1590 CAAGGACGGGCACATTAAAGATCAAGACTTTCGGGCTGTGCAAGGAGGGATCAAGGACGG 1649

Qy 609 CCGGTCAAGCGGTGAGCGGACCATCCCTTTACACGGCGCTGAGTGTGCCAGGCGGG 668
Db 1650 TGCCACCATGAAGACCTTTTGGCGCACACTGTGAGTACCTGGCCCCCGAGGTGTGGAGGA 1709

Qy 669 CCGCGCCAGCGGTGGCGGTGGAACAGCGCGTGAAGTGTGGGCGCTTCGCGGTGCTCAT 728
Db 1710 CAATGACTACGGCGCTGAGTGAAGCTTGTGGGCGCTGGGCGTGTGATGACGAGATGAT 1769
```

Db 1134 GAGGACTGTACGTCTTTGCCCAGAGTACGCACCTGCTGGGACCTGTTTGATCATCATC 1193
Qy 421 CCTCCACAGTGGGCTCCCTGAGGACACCGGTGAAGCGCTGTGTGACGACGTGGGCTG 480
Db 1194 CCTCCACAGTGGGCTCCCTGAGGACACCGGTGAAGCGCTGTGTGACGACGTGGGCTG 1253
Qy 481 GCGGTGACCTTCAATGACCGGCGGACGTGTGTGACCGGACATCAAGCCGAGAACGTG 540
Db 1254 GCGGTGACCTTCAATGACCGGCGGACGTGTGTGACCGGACATCAAGCCGAGAACGTG 1313
Qy 541 CTGCTGTTCACCGCGAGTGGCGCGGTAAAGTGGCCGACCTTCGACATGACGCGCGC 600
Db 1314 CTGCTGTTCACCGCGAGTGGCGCGGTAAAGTGGCCGACCTTCGACATGACGCGCGC 1373
Qy 601 GTGGGTGCGCGGTCAAGCGGTGAGCGGACACATCCCTTACACGCGCGCTGAGGTGTC 660
Db 1374 GTGGGTGCGCGGTCAAGCGGTGAGCGGACACATCCCTTACACGCGCGCTGAGGTGTC 1433
Qy 661 CAGCGCGCGCGCGCGGCTGGCGGTGGAACACGCGGTGAGCTGTGGGCTTCGCG 720
Db 1434 CAGCGCGCGCGCGCGGCTGGCGGTGGAACACGCGGTGAGCTGTGGGCTTCGCG 1493
Qy 721 GTGCTCATCTTCTGCGTGTCAACCGGCAACTTCCGCTGGGAGGCGGTGCGGCGCGAC 780
Db 1494 GTGCTCATCTTCTGCGTGTCAACCGGCAACTTCCGCTGGGAGGCGGTGCGGCGCGAC 1553
Qy 781 GCCTTCTTCAGAGGTTCGTGCGTGGCAGCGGCGCGCTTCGCGGCGCTTCGCGAG 840
Db 1554 GCCTTCTTCAGAGGTTCGTGCGTGGCAGCGGCGCGCTTCGCGGCGCTTCGCGAG 1613
Qy 841 TGGCGCGCTTACCGAGCGCGCTGCGGATGTTCCAGCGCTTACCTGAGCGCTGAGCGCC 900
Db 1614 TGGCGCGCTTACCGAGCGCGCTGCGGATGTTCCAGCGCTTACCTGAGCGCTGAGCGCC 1673
Qy 901 GAGCGCGCGCGCGCGAGGAGTGTTCGCTTCTCAAGCAGCAGCTCACGTCGCGAG 960
Db 1674 GAGCGCGCGCGCGCGAGGAGTGTTCGCTTCTCAAGCAGCAGCTCACGTCGCGAG 1733
Qy 961 CTGCGCGCGCGCGCTTCGACCGCGCGCAAGCCCGCGGACCGCGCGCGCGCGCGCC 1020
Db 1734 CTGCGCGCGCGCGCTTCGACCGCGCGCAAGCCCGCGGACCGCGCGCGCGCGCGCC 1793
Qy 1021 GGGCCACTGCGCTGAGGCGCGCTGGCGCTCAAGCGGACGCTGACGAGGCGC 1080
Db 1794 GGGCCACTGCGCTGAGGCGCGCTGGCGCTCAAGCGGACGCTGACGAGGCGC 1853
Qy 1081 AGCGCTCCCGCGCGCGCGCGCGCGCTGGCGCTCAAGCGGACGCTGACGAGGCGC 1140
Db 1854 AGCGCTCCCGCGCGCGCGCGCGCGCTGGCGCTCAAGCGGACGCTGACGAGGCGC 1913
Qy 1141 GTGCCAGTGGCGCGTGGCGTTCCTGTCGCGAGCGCGCGCTAGCTCCCGAGGCGCGCGC 1200
Db 1914 GTGCCAGTGGCGCGTGGCGTTCCTGTCGCGAGCGCGCGCTAGCTCCCGAGGCGCGCGC 1973
Qy 1201 GCGCGGACCGCGCGCGCGGACGAGGACGAGGAGGAGGAGTGTGCTGGCCACGCGCATC 1260
Db 1974 GCGCGGACCGCGCGCGCGGACGAGGACGAGGAGGAGGAGTGTGCTGGCCACGCGCATC 2033
Qy 1261 GAGATCTGCGTCTGA 1275
Db 2034 GAGATCTGCGTCTGA 2048

RESULT 2

US-09-816-094-3
; Sequence 3, Application US/09816094
; Patent No. 6534299
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536-CIP

; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: Human
US-09-816-094-3

Query Match 66.6%; Score 849.6; DB 4; Length 7301;
Best Local Similarity 99.5%; Pred. No. 5.2e-151;
Matches 852; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 420 CCCTCCCCAGGTGGGCTCCCTGAGGACACGCTGTAAGCGCTGTGTGACGACGTGGGCT 479
Db 3448 CCGCCCGCAGGTGGGCTCCCTGAGGACACGCTGTAAGCGCTGTGTGACGACGTGGGCT 3507
Qy 480 GGGCTGGACTTCATGACGCGGCGGACGCTGGTGACCGGACATCAAGCCGAGAACGT 539
Db 3508 GGGCTGGACTTCATGACGCGGCGGACGCTGGTGACCGGACATCAAGCCGAGAACGT 3567
Qy 540 GCTGCTGTTGACCGCGAGTGGCGCGGTAAAGCTGGCGGACTTCGGCATGACGCGCG 599
Db 3568 GCTGCTGTTGACCGCGAGTGGCGCGGTAAAGCTGGCGGACTTCGGCATGACGCGCG 3627
Qy 600 CGTGGGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 659
Db 3628 CGTGGGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 3687
Qy 660 CAGCGCGCGCGCGCGACGCGGCTGGCGGTGGACGCGGCGTGGGCGCTTCGG 719
Db 3688 CAGCGCGCGCGCGCGACGCGGCTGGCGGTGGACGCGGCGTGGGCGCTTCGG 3747
Qy 720 CGTGCTCATCTTCTGCGTGTCTACCGGCACTTCCGCTGGGAGCGCGCTCGGCGCGCA 779
Db 3748 CGTGCTCATCTTCTGCGTGTCTACCGGCACTTCCGCTGGGAGCGCGCTCGGCGCGCA 3807
Qy 780 CGCTTCTTCGAGGAGTGTGTGCTGCGTGGCAGCGGCGCGCTTCGCGGCGTGTGCTTGC 839
Db 3808 CGCTTCTTCGAGGAGTGTGTGCTGCGTGGCAGCGGCGCGCTTCGCGGCGTGTGCTTGC 3867
Qy 840 GTGGCGCGCTTCAACGAGCGCGCGCTGGCATGTTCCAGCGCTTACTGSCCTTGAGCGC 899
Db 3868 GTGGCGCGCTTCAACGAGCGCGCGCTGGCATGTTCCAGCGCTTACTGSCCTTGAGCGC 3927
Qy 900 CAGCGCGCGCGCGCGACGAGAGTGTTCGCTTCTCAAGCAGAGCTCACGCTCCGA 959
Db 3928 CAGCGCGCGCGCGCGACGAGAGTGTTCGCTTCTCAAGCAGAGCTCACGCTCCGA 3987
Qy 960 GCTGCGCGCGCGCGCTTCGACCGCGCGCGCAAGCCCGCGGACCGCGCGCGCGCGC 1019
Db 3988 GCTGCGCGCGCGCGCTTCGACCGCGCGCGCAAGCCCGCGGACCGCGCGCGCGCGC 4047
Qy 1020 CCGGCGCACTCGGCTCGAGCGCGCTGGCGCTCAAGCGACGCTGCTGACCGAGCGG 1079
Db 4048 CCGGCGCACTCGGCTCGAGCGCGCTGGCGCTCAAGCGACGCTGCTGACCGAGCGG 4107
Qy 1080 CAGCGGCTCCCGCGCGCGCGCGCGCTGCGGTGCGTGTGCGCTTGCCTGCGCGTGC 1139
Db 4108 CAGCGGCTCCCGCGCGCGCGCGCGCTGCGGTGCGTGTGCGCTTGCCTGCGCGTGC 4167
Qy 1140 GTTGCCAGTGTGCGCGTGTGCGCGCGCGCGCTGCGGTGCGTGTGCGCGCGCGCGC 1199
Db 4168 GTTGCCAGTGTGCGCGTGTGCGCGCGCGCGCTGCGGTGCGTGTGCGCGCGCGCGC 4227
Qy 1200 CCGCGGACCGAGCGCGCGCGGACGAGGAGGAGGAGGAGTGTGCTGCGCGCGCAT 1259
Db 4228 CCGCGGACCGAGCGCGCGCGGACGAGGAGGAGGAGGAGTGTGCTGCGCGCGCAT 4287
Qy 1260 CGAGATCTGCGTCTGA 1275
Db 4288 CGAGATCTGCGTCTGA 4303

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 15:10:03 ; Search time 76 Seconds
(without alignments)
7404.783 Million cell updates/sec

Title: US-10-016-985-1

Perfect score: 1275
Sequence: 1 atgagctggctgccaga.....ccatcgagatctgcgtctga 1275

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/pctus_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1275	100.0	2598	4	US-09-816-094-1
2	849.6	66.6	7301	4	US-09-816-094-3
3	97.6	7.7	2181	4	US-09-417-137-70
4	97.6	7.7	2184	4	US-09-417-137-138
5	97.6	7.7	2610	2	US-09-212-771-1
6	97.6	7.7	2610	3	US-09-091-058-1
7	93.4	7.3	1599	3	US-09-256-465-1
8	93.4	7.3	1599	4	US-09-167-322-3
9	82.2	6.4	5228	4	US-09-428-711A-15
10	75.4	5.9	1485	4	US-09-801-876B-1
11	74.4	5.8	1008	4	US-09-394-453-3
12	74.4	5.8	2549	3	US-09-467-082-3
13	74.4	5.8	2549	4	US-09-394-455-5
14	74.4	5.8	2608	4	US-09-394-455-35
15	73.8	5.8	2146	4	US-09-842-307-1
16	72.6	5.7	2902	4	US-09-579-664B-4
17	72.4	5.7	1282	2	US-08-878-989-12
18	72.4	5.7	1282	3	US-09-272-796-12
19	72.4	5.7	1282	4	US-09-016-434-953
20	72	5.6	1619	4	US-09-394-455-14
21	69.6	5.5	2746	2	US-09-016-000-12
22	69	5.4	44377	2	US-08-804-227C-7
23	69	5.4	44377	2	US-08-804-198-1
24	67.8	5.3	722	3	US-08-998-416-800
25	67.8	5.3	2673	3	US-09-428-711A-1
26	67.8	5.3	12001	1	US-08-458-568A-11
27	67.6	5.3	4257	2	US-08-690-473-1

```

28 67.6 5.3 4257 3 US-09-253-821A-1 Sequence 1, Appli
29 67.6 5.3 4257 3 US-08-843-659-1 Sequence 1, Appli
30 67.4 5.3 2908 4 US-09-930-181-1 Sequence 1, Appli
31 66.6 5.2 5207 4 US-09-853-664A-1 Sequence 1, Appli
32 66.2 5.2 3995 4 US-09-423-711A-13 Sequence 13, Appli
33 66.2 5.2 441529 3 US-09-103-840A-1 Sequence 1, Appli
34 64.9 5.1 3228 4 US-09-579-664B-5 Sequence 5, Appli
35 64.6 5.1 1965 4 US-09-252-991A-9230 Sequence 9230, Ap
36 64.6 5.1 2805 4 US-09-252-991A-9064 Sequence 9064, Ap
37 64.6 5.1 3364 4 US-09-930-181-3 Sequence 3, Appli
38 64.4 5.1 2132 2 US-09-159-385-3 Sequence 3, Appli
39 64.4 5.1 2132 3 US-09-186-277-3 Sequence 3, Appli
40 64.4 5.1 4403785 3 US-09-103-840A-2 Sequence 2, Appli
41 64 5.0 1788 4 US-09-417-197-68 Sequence 68, Appli
42 64 5.0 2211 4 US-09-354-455-39 Sequence 39, Appli
43 62.6 4.9 999 4 US-09-252-991A-5206 Sequence 5206, Ap
44 62.6 4.9 1710 4 US-09-252-991A-5271 Sequence 5271, Ap
45 62.4 4.9 3471 2 US-08-715-568A-2 Sequence 2, Appli

```

ALIGNMENTS

RESULT 1
US-09-816-094-1
; Sequence 1, Application US/09816094
; Patent No. 6534299
; GENERAL INFORMATION:

APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CLO00536-CIP
CURRENT APPLICATION NUMBER: US/09/816,094
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2598
TYPE: DNA
ORGANISM: Human
US-09-816-094-1

Query Match 100.0%; Score 1275; DB 4; Length 2598;

Best Local Similarity 100.0%; Pred. No. 8.3e-231;

Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 ATGAGCGTGGCTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGCGCGGG 60
Db 774 ATGAGCGTGGCTGCCAGAGCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGCGCGGG 833
Qy 61 ACTGCCCTGGCGCTGGTGGCGGTGTCCTTCTCACTGAAGACATGCGAGCGCTGACT 120
Db 834 ACTGCCCTGGCGCTGGTGGCGGTGTCCTTCTCACTGAAGACATGCGAGCGCTGACT 893
Qy 121 CTCGCGACATGCGCGCGCGAGCGCTGAGCGCGCGCGCTGAGCGCGCGCTGAGCGCGCTGACT 180
Db 894 CTCGCGACATGCGCGCGCGAGCGCTGAGCGCGCGCGCTGAGCGCGCGCTGAGCGCGCTGACT 953
Qy 181 AAAGGCACTTATGGGAGGTGACCTGGTGTCTACAGGCGCAGCAGCAAAAATGCGCA 240
Db 954 AAAGGCACTTATGGGAGGTGACCTGGTGTCTACAGGCGCAGCAGCAAAAATGCGCA 1013
Qy 241 CTGAAGTTTGTGAACAGAGCAAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 300
Db 1014 CTGAAGTTTGTGAACAGAGCAAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 1073
Qy 301 ACCAAGCTTCTTCTTCCAGCGCGCTTCTCATCAAGGTCTTTGACGTGTTTGTAGACA 360
Db 1074 ACCAAGCTTCTTCTTCCAGCGCGCTTCTCATCAAGGTCTTTGACGTGTTTGTAGACA 1133
Qy 361 GAGGACTGCTACGCTTTTGTCCAGGAGTACGACCTGCTGGGACCTGTTTGACATCATC 420

```


Db 290 CAGAAAGCACACCCCTGGCACTGAGCAGCTCCCGAAACCCCGCAGCTCCCTCGGTGGC 349
Qy 280 TTCCTACGGGAGGTAGCATCACCAACAGCTCTCTCCACGCCCTTCATCATCAAGATC 339
Db 350 TTCCTGTACGAGTCTGTGTGGGCTCTCGCTGGCGCGCACTCAGCCATGTCAGCGCC 409
Qy 340 TTTGACGTGTCCTTTGACAGAGGACTGCTACGCTCTTTGCCAGGAGTAGCAGCTGCT 399
Db 410 TACGGCATGTCATCGATCGGACACTCTTACAGCTTCTTCGACGGAGCCGCTCTGCAC 469
Qy 400 GGGGACCTGTTTACATCATCCCTCCCGAGGTGGGCTCCCTGAGGACACAGCTGAAGCGC 459
Db 470 GGGGACCTCATGGCTTCATCCAGCCCAAGTGGGCTCCCGCAGCCCGGCTGCACCGC 529
Qy 460 TGTGTGACGAGCTGGGCTGGCTGGCTGACTTATGACAGGGGGGAGCTGCTGTCACCGC 519
Db 530 TGGCGCCGCCAGCTGGGCTCGGCTCGGCTGAGTATCATCCAGCCCGCGGCTGTATCCGG 589
Qy 520 GACATCAAGCCCGAGAGCTGCTGCTGTTCGACCGGAGTCCCGCGAGTCCCGCGTAAGCTGGC 579
Db 590 GACTGAGCCGGAGAGCTGCTGCTGTGCGACCCCGGCTCCCGGCTTCAAGCTGACC 649
Qy 580 GACTTCGGCATGACGCGCGGCTGGGCTGCGCGCTCAAGCGCTGAGCGGACCATCCCT 639
Db 650 GACTTCGGCATGACGCGCGGCTGCGCGGAGCTGCTGCGCTGGCGGCGGCGCCATCCCC 709
Qy 640 TACAGGGGCTGAGGTGTC---CAGCGGGGCGCGCGGCTGGCGGTGGAGTGGAGCAG 696
Db 710 TACAGCGCGCCGAGCTGTGGCGCCCGCGCGCTCCCGAGGCGCTGCCCATTCAGCGCC 769
Qy 697 GGCCTGAGAGCTGTGGGCTTCGGGCTGCTCATCTTCTGCTGCTGCTCAGCGCACTTCCCG 756
Db 770 GCCCTGAGAGCTGGGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
Qy 757 TGGGA---GGCGGCTGGGCGCGCGAGCTTCTTCGAGGAGTTCGTGCTGCGTGGCAGCG 813
Db 830 TGGGACCGGCGCTGGCGAGCGAGCGAGCGCTTCTACGAGGACTTCTCATCTGCGAGCG 889
Qy 814 GCGCGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
Db 890 TCGGGCGAGCGCGGCGAGCGCGCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949
Qy 874 TTCAGAGCTTACTG 888
Db 950 CTTCGCGGGGCTG 964

RESULT 14

US-10-288-798-40

; Sequence 40, Application US/10288798

; Publication No. US20030207299A1

; GENERAL INFORMATION:

; APPLICANT: BANDMAN, Olga; NGUYEN, Dannel B.;
; APPLICANT: WALIA, Narender K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Valda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 7477062CB1
US-10-288-798-40

Query Match 18.4%; Score 235; DB 12; Length 1156;

Best Local Similarity 57.7%; Pred. No. 4,7e-46;

Matches 459; Conservative 0; Mismatches 330; Indels 6; Gaps 2;

Qy 100 GAGACATGAGGCGCTGACTCTCCGCACACTGCGCCAGCGAGCTCACCAGCACTAC 159
Db 236 GAGACATGATGAGCTGAGTGTCTCAGCCCTGTGCGAGCCGAGGTGGAGCTTAC 295
Qy 160 GAACTAGTCCGGAGCTGGCAAGCACTATGGGAAGTTGACCTGGTGGTCTACAAG 219
Db 296 GAGGAAGTGTCCCTGGCCAGGGTCTGATGGCGCGTCTCTTGTTCACCCATCGT 355
Qy 220 GGCACAGGCAAAATGGCACTGAAGTTGTGAACAGAGCAAAACCAAGCTGAGAAC 279
Db 356 CAGAAAGGCACACCCCTGGCACTGAAGCACTCCCGAAACCCCGCAGCTCCCTCGTGGC 415
Qy 280 TTCCTACGGAGGTGAGCATCACCAAGCGCTCTCTCCAGCGCCCTTCATCATCAAGTC 339
Db 416 TTCCTGTACAGTTCTGTGTGGGCTCTCGTGGGCGCGCACTCAGCCATCGTGACGGC 475
Qy 340 TTTGACGTGCTTTTGAGACAGAGGACTGCTAGTCTTTTGGCCAGGAGTACGACCTGCT 399
Db 476 TACGGCATTTGGCATCGAGTCGGCACACTCTCTACAGCTTCTGACGGAGCCGCTCTGCAC 535
Qy 400 GGGGACCTGTTGACATCATCCCTCCCGAGTGGGCTCTCTGAGGACACGGTGAAGCGC 459
Db 536 GGGGACCTCATGGCTTTCATCCAGCCCAAGTGGGCTCTCCCGCAGCGCGGTCACCGC 595
Qy 460 TGTGTGACGAGCTGGGCTTGGGCTGAGACTTATGACAGGGGCGGAGCTGGTGCACCGC 519
Db 596 TGGCGCCCGAGCTGGCTCCGCTCGGCTGGAGTATCATCACGCCCGCGCTGTGTACCGG 655
Qy 520 GACATCAAGCCCGAGAACGTGCTGCTGTTCGACCGGAGTTCGCCCGCGCTAAAGTGGCC 579
Db 656 GACCTGAAGCCGGAGAACGTCTGTGTGGAGCCCGGCTTCCCGGCTTCAAGTGAAC 715
Qy 580 GACTTCGGCATGAGCGCGCGCTGGGCTGCGGCTGAGCGGTGAGCGGACCATCCCT 639
Db 716 GACTTCGGCCACAGAGGCTCGCGGAGCGCTGCGCTGCGCTGCGCGCGCGCCATCCCC 775
Qy 640 TACACGGGCTGAGGTGTC---CAGGGGGCGCGCGCGCGCTGGCGGTGGAGCAG 696
Db 776 TACAGGCCCCGAGCTCTCGGCGCCCCCGCGCTCCCGAGGCGCTGCCCATTCAGCCC 835
Qy 697 GCGGTGGAAGCTGGGCTTTCGGCGTGTCTCATCTTCTGGGTGCTCAGCGGCAACTTCCCG 756
Db 836 GCGCTGGAAGCTGGGCTGGGCTGCTGCTCTTCTGCTGCTCTCTCAGGCGCTACTTCCCC 895

```
; TITLE OF INVENTION: METHYLATED CpG ISLANDS
; FILE REFERENCE: 01107.00128
; CURRENT APPLICATION NUMBER: US/09/861,893
; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/206,158
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/206,161
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 788
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-861-893-38

Query Match      60.2%; Score 767.6; DB 9; Length 788;
Best Local Similarity 99.4%; Pred. No. 8.9e-172;
Matches 781; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 420 CCTCTCCCCAGGTGGGGTCCCTGAGGACACGGTGAAGCGTGTGTGTCAGCAGCTGGGCGCT 479
Db 785 CCGCCCGCAGGTGGGGTCCCTGAGGACACGGTGAAGCGTGTGTGTCAGCAGCTGGGCGCT 726
Qy 480 GGGCTGGACTTCATGACCGGGCGGCGAGCTGGTGCACGGGACATCAAGCCCGAGACGT 539
Db 725 GGGCTGGACTTCATGACCGGGCGGCGAGCTGGTGCACGGGACATCAAGCCCGAGACGT 666
Qy 540 GCTCTGTTTCAGCCGAGTGGCCCGCGTAAAGCTGGCGACATTCGGCATGACGCCCG 599
Db 565 GCTGCTGTTTCAGCCGAGTGGCCCGCGTAAAGCTGGCGACATTCGGCATGACGCCCG 606
Qy 600 CGTGGGTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 659
Db 605 CGTGGGTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 546
Qy 660 CCAGCGGGCGCGCCGACCGGGTGGCGGTGGACACGGGCGTGAAGCTGGGCGCTTCGG 719
Db 545 CCAGCGGGCGCGCCGACCGGGTGGCGGTGGACACGGGCGTGAAGCTGGGCGCTTCGG 486
Qy 720 CGTCTCATCTTTCGCGTGTCTACCGGCACTTCCTCGTGGAGCGCGCTGCGGCGCGCA 779
Db 485 CGTCTCATCTTTCGCGTGTCTACCGGCACTTCCTCGTGGAGCGCGCTGCGGCGCGCA 427
Qy 780 CGCCTTCTTCAGAGGATTCGCTGCGCTGCGACCGGGCGCGCTGCGGGGCTGCTTCGCA 839
Db 426 CGCCTTCTTCAGAGGATTCGCTGCGCTGCGACCGGGCGCGCTGCGGGGCTGCTTCGCA 367
Qy 840 GTGGCGCGGCTTCACGAGCGCGCGCTGCGCATGTTCCAGGGCTTACTGGCCTTGGAGCC 899
Db 366 GTGGCGCGGCTTCACGAGCGCGCGCTGCGCATGTTCCAGGGCTTACTGGCCTTGGAGCC 307
Qy 900 CGAGCGCGCGCGCCCGCAGCCAGGAGGTGTTCCGCTTCTTCAAGCACGAGCTCAGTCCGA 959
Db 306 CGAGCGCGCGCGCCCGCAGCCAGGAGGTGTTCCGCTTCTTCAAGCACGAGCTCAGTCCGA 247
Qy 960 GCTGCGCGCGCGCGCTTCGACACCGCGCGCGCAAGCCCGCGGGGACCGCGCGCGCGCG 1019
Db 246 GCTGCGCGCGCGCGCTTCGACACCGCGCGCGCAAGCCCGCGGGGACCGCGCGCGCGCG 187
Qy 1020 CGGGCCACTGCGCTTCGAGGCGCGCTGGCGCGCTCAAGGAGCGTGTGACCGAGAGCGG 1079
Db 186 CGGGCCACTGCGCTTCGAGGCGCGCTGGCGCGCTCAAGGAGCGTGTGACCGAGAGCGG 127
Qy 1080 CAGCGGCTTCCCGCGCGCGCGCGCGCGCGTGGGTGCGGTGCGCTTGGCGCGCGGTGCC 1139
Db 126 CAGCGGCTTCCCGCGCGCGCGCGCGCGCGTGGGTGCGGTGCGCTTGGCGCGCGGTGCC 67
Qy 1140 GGTGCGAGTGCCTGCGCGGTGCTGTGCGCGCGCGCGCGCGGTAGTCCCGAGGGCGCGCC 1199
Db 66 GGTGCGAGTGCCTGCGCGGTGCTGTGCGCGCGCGCGCGCGGTAGTCCCGAGGGCGCGCC 7
Qy 1200 CGGCGG 1205
```

```
Db 6 CGGCGG 1
```

RESULT 13

```
US-10-353-690-49
; Sequence 49, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodrigue-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MPI02-018P1RNMNIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 1247
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-353-690-49
```

```
Query Match      18.6%; Score 236.6; DB 12; Length 1247;
Best Local Similarity 57.9%; Pred. No. 2e-46;
Matches 460; Conservative 0; Mismatches 329; Indels 6; Gaps 2;

Qy 100 GAAGACATGAGCCCTGACTCTCCGACACACTGCGCGCGCAGCGACGCTCACCAGCACTAC 159
Db 170 GAGSACATGATGACGCTGAGTGTCTAGACCTGCTCCGAGCGAGGTGGACGAGCTCTAC 229
Qy 160 GAAGTACTCGGAGCTGGCAAGGCACTATGGAAGTTGACCTGTTGGTGTCTACAAG 219
Db 230 GAGGAAGTGGTCCCTGGGCGGAGGTGCTATGCGCGGCTCTCTGTTGTCACCCATCGT 289
Qy 220 GGCACAGGCAAAATGGCACTGAAATTTGTGCAACAGAGCAAAACCAAGCTGAAGAAC 279
```

Db	3568	GCTGCTTTT	CGAACCGAGTGC	CGCCGCTAAAGCTG	CGCGCATTTG	CGCATGACGCGCG	3627
Qy	600	CTGGGGTGC	CGCGTCAAGCGCGT	AGCGGCACCATCC	TTACACGCGCGCT	GAGGTGTG	659
Db	3628	CTGGGGTGC	CGCGTCAAGCGCGT	AGCGGCACCATCC	TTACACGCGCGCT	GAGGTGTG	3687
Qy	660	CGAGCGGGG	CGCGCGACGCGGCT	GGCGGTGGAGACG	GGCGTGGACGTGG	GCCTTCG	719
Db	3688	CGAGCGGGG	CGCGCGACGCGGCT	GGCGGTGGAGACG	GGCGTGGACGTGG	GCCTTCG	3747
Qy	720	CGTGTCTCAT	CTTTGCGTGTCT	CACCGGCAACTTCC	CGTGGAGCGCGCT	CGGCGCCGA	779
Db	3748	CGTGTCTCAT	CTTTGCGTGTCT	CACCGGCAACTTCC	CGTGGAGCGCGCT	CGGCGCCGA	3807
Qy	780	CGCCTTCTT	CGAGAGTTCGTG	CGCTGCGACGCGG	CGCGCTCGCGGG	CTGCTTCGCA	839
Db	3808	CGCCTTCTT	CGAGAGTTCGTG	CGCTGCGACGCGG	CGCGCTCGCGGG	CTGCTTCGCA	3867
Qy	840	GTGGCGCGCT	TACCGAGCCGCGCT	GGCGATGTTTCAC	GGCTTACTGGCC	CTGAGCC	899
Db	3868	GTGGCGCGCT	TACCGAGCCGCGCT	GGCGATGTTTCAC	GGCTTACTGGCC	CTGAGCC	3927
Qy	900	CGAGCGCGCG	CGCCACGCAAGGAGT	GTTCGGCTTCTCAA	GACACAGACTCAC	GTCCTCGA	959
Db	3928	CGAGCGCGCG	CGCCACGCAAGGAGT	GTTCGGCTTCTCAA	GACACAGACTCAC	GTCCTCGA	3987
Qy	960	GCTGCGCGCG	CGGCGCTTGCAC	CGCGCGCGAAGCC	CCCCCGGGGA	CGCCCGCGCCCGC	1019
Db	3988	GCTGCGCGCG	CGGCGCTTGCAC	CGCGCGCGAAGCC	CCCCCGGGGA	CGCCCGCGCCCGC	4047
Qy	1020	CGGGCCACT	CGCGCTCGAGCGCT	TGGCGCTCCTAC	GCGACGGTGTGTG	ACCGAGACGG	1079
Db	4048	CGGGCCACT	CGCGCTCGAGCGCT	TGGCGCTCCTAC	GCGACGGTGTGTG	ACCGAGACGG	4107
Qy	1080	CAGCGGCTC	CGGCGCCCGCGCG	CCCCCGCGTGGG	TCGGCTTGCCCT	TGCGCGCGGTGCC	1139
Db	4108	CAGCGGCTC	CGGCGCCCGCGCG	CCCCCGCGTGGG	TCGGCTTGCCCT	TGCGCGCGGTGCC	4167
Qy	1140	GGTGCCAGT	GTCCGTGCTGTG	CCGAGCCCGCGCT	AGCTCCCCAG	GGGGCCCC	1199
Db	4168	GGTGCCAGT	GTCCGTGCTGTG	CCGAGCCCGCGCT	AGCTCCCCAG	GGGGCCCC	4227
Qy	1200	CGGCGGAC	CGACGGCGCGGAC	ACAGACAAAGG	ACAGTGGTGTG	GTGSCACGCCAT	1259
Db	4228	CGGCGGAC	CGACGGCGCGGAC	ACAGACAAAGG	ACAGTGGTGTG	GTGSCACGCCAT	4287
Qy	1260	CGAGATCT	CGGTCTGA	1275			
Db	4288	CGAGATCT	CGGTCTGA	4303			

```

RESULT 11
US-10-233-613-3
; Sequence 3, Application US/10233613
; Publication No. US20030022339A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000536DIV
; CURRENT APPLICATION NUMBER: US/10/233,613
; CURRENT FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-233-613-3

```

Query Match	Score 849.6;	DB 14;	Length 7301;
66.6%;			

Best Local Similarity		99.5%	Pred. No. 4.9e-191;						
Matches		852;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps
Qy	420	CCCTCCCCAGGTGGGGCTCCCTGAGGACACGGGTGAAGCGCTGTGTGCAGCAGCTGTGGCGCT	479						
Db	3448	CCGCCCGCAGGTGGGGCTCCCTGAGGACACGGGTGAAGCGCTGTGTGCAGCAGCTGTGGCGCT	3507						
Qy	480	GGCGCTGGACTTATATGCAACGGGGCGGCACTGTGTGCAACGGGACATCAAGAGCCCGAGAACGT	539						
Db	3508	GGCGCTGGACTTATATGCAACGGGGCGGCACTGTGTGCAACGGGACATCAAGAGCCCGAGAACGT	3567						
Qy	540	GCTGCTCTTCGACCGCGAGTCCGCCCGGTAAAGCTGGCCGCACTTCGGCATGACGCCCGC	599						
Db	3568	GCTGCTCTTCGACCGCGAGTCCGCCCGGTAAAGCTGGCCGCACTTCGGCATGACGCCCGC	3627						
Qy	600	CGTGGGCTGCCGGCTCAAGCGCGGTGACGGCACCATCCCTTACACGGCGGCTGAGGTGTG	655						
Db	3628	CGTGGGCTGCCGGCTCAAGCGCGGTGACGGCACCATCCCTTACACGGCGGCTGAGGTGTG	3687						
Qy	660	CCAGCGGGCGCGCGCAACGGGCTGGCGGTGAGACGGGCGTGGAACGTGTGGCGCTTCGG	719						
Db	3688	CCAGCGGGCGCGCGCAACGGGCTGGCGGTGAGACGGGCGTGGAACGTGTGGCGCTTCGG	3747						
Qy	720	CGTGTCTCATCTTCTGCGTGTCAACGGCAACTTCCCGTGGAGGCGGCGCTCGGGCGCCGA	779						
Db	3748	CGTGTCTCATCTTCTGCGTGTCAACGGCAACTTCCCGTGGAGGCGGCGCTCGGGCGCCGA	3807						
Qy	780	CGCGCTTCTCGAGGAGTTCGTGGCTGGCAAGCGGGCGCGCTCCCGGGGTGCTTCGCA	839						
Db	3808	CGCGCTTCTCGAGGAGTTCGTGGCTGGCAAGCGGGCGCGCTCCCGGGGTGCTTCGCA	3867						
Qy	840	GTGGCGCGCTTCAACGAGCCGCGCTGCGCATGTTCCAGGGCTTACTGCGCCCTGGAGCC	899						
Db	3868	GTGGCGCGCTTCAACGAGCCGCGCTGCGCATGTTCCAGGGCTTACTGCGCCCTGGAGCC	3927						
Qy	900	CGAGCGCGCGGGCCCGACCAAGGAGTGTTCGCTTCCTCAAGCACAGACTCAGTCCGA	959						
Db	3928	CGAGCGCGCGGGCCCGACCAAGGAGTGTTCGCTTCCTCAAGCACAGACTCAGTCCGA	3987						
Qy	960	GCTGCGCGCGCGGCGCTTCGCAACGGCGCGGAGACCGCCCGCGGGAACCGCGCGCCCGCGC	1019						
Db	3988	GCTGCGCGCGCGGCGCTTCGCAACGGCGCGGAGACCGCCCGCGGGAACCGCGCGCCCGCGC	4047						
Qy	1020	CGGGCCACTGCGCTTCGAGCGGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGG	1079						
Db	4048	CGGGCCACTGCGCTTCGAGCGGCTGGCGCGCTCAAGCGGACGCTGCTGACCGAGAGCGG	4107						
Qy	1080	CAGCGGTCCCGGCCCGCGCCCGCGCTCGGGTCGGTCCCTTGCCCGTGGCGGTGCC	1139						
Db	4108	CAGCGGTCCCGGCCCGCGCCCGCGCTCGGGTCGGTCCCTTGCCCGTGGCGGTGCC	4167						
Qy	1140	GGTGCAAGTGGCGGTGGCGGTGCTGTGCGGAGCCCGGCTAGCTCCCGAGGGGCCCGC	1199						
Db	4168	GGTGCCAGTGGCGGTGGCGGTGCTGTGCGGAGCCCGGCTAGCTCCCGAGGGGCCCGC	4227						
Qy	1200	CGGCGGACCGACGGCGCGCGGACAAAGACAAAGGGCAGGTGGTGTGGCCACGGCCAT	1259						
Db	4228	CGGCGGACCGACGGCGCGCGGACAAAGACAAAGGGCAGGTGGTGTGGCCACGGCCAT	4287						
Qy	1260	CGAGATCTGCGTCTGA	1275						
Db	4288	CGAGATCTGCGTCTGA	4303						

```

RESULT 12
US-09-861-893-38/c
; Sequence 38, Application US/09861893
; Patent No. US20020045257A1
; GENERAL INFORMATION:
; APPLICANT: Feinberg, Andrew
; APPLICANT: Strichman-Almashanu, Liora
; APPLICANT: Jiang, Shan
; TITLE OF INVENTION: METHODS FOR ASSAYING GENE IMPRINTING AND

```

```
QY 955 TCCAGAGTCCGCGCGCGGCTCCGACCGCGCGCGAAGCCCCCGGGGACGCCCGCGCC 1014
Db 721 TCCAGAGTCCGCGCGCGGCTCCGACCGCGCGCGAAGCCCCCGGGGACGCCCGCGCC 780
QY 1015 GCCCGCGCGGCACTGCGGCTCAGAGCGCTGGCGCGCTCAAGCGAGCGGTGCTGACCGAG 1074
Db 781 GCCCGCGCGGCACTGCGGCTCAGAGCGCTGGCGCGCTCAAGCGAGCGGTGCTGACCGAG 840
QY 1075 AGCGGAGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1134
Db 841 AGCGGAGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 1135 GTCCGCGTGCAGTGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1194
Db 901 GTCCGCGTGCAGTGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 1195 CCCCCGCGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1254
Db 961 CCCCCGCGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1255 GCCATCGAGATCTCGTCTGA 1275
Db 1021 GCCATCGAGATCTCGTCTGA 1041

RESULT 9
US-09-816-094-3
; Sequence 3, Application US/09816094
; Patent No. US2002064851A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000536-CIP
; CURRENT APPLICATION NUMBER: US/09/816,094
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: Human
US-09-816-094-3

Query Match 66.6%; Score 849.6; DB 9; Length 7301;
Best Local Similarity 99.5%; Pred. No. 4.9e-191;
Matches 852; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 420 CCCTCCCGAGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCT 479
Db 3448 CCGCCCGAGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCT 3507
QY 480 GCGCTGGAATTCATGACGCGCGCGAGCTGTGTGACCGCGCATCAAGCCCGAGAACGT 539
Db 3508 GCGCTGGAATTCATGACGCGCGCGAGCTGTGTGACCGCGCATCAAGCCCGAGAACGT 3567
QY 540 GCTGCTGTGACCGCGAGTGCCTGTAAGCTGCGCGCATTCGCGCATGACGCGCG 599
Db 3568 GCTGCTGTGACCGCGAGTGCCTGTAAGCTGCGCGCATTCGCGCATGACGCGCG 3627
QY 600 GTCGGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 659
Db 3628 GTCGGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 3687
QY 660 CCAGCGCGCGCGCGCGCGCTGGCGGTGACACGCGCGCTGTCAGCTGTGGGCTTCGG 719
Db 3688 CCAGCGCGCGCGCGCGCGCTGGCGGTGACACGCGCGCTGTCAGCTGTGGGCTTCGG 3747
QY 720 CGTGTCTATCTTCTCGTGTCTACCGGCAATCTTCCGTTGGAGGCGCGCTGGCGCGCG 779
Db 3748 CGTGTCTATCTTCTCGTGTCTACCGGCAATCTTCCGTTGGAGGCGCGCTGGCGCGCG 3807
```

```
QY 780 CCGCTTCTTCGAGGAGTTCGTGCGCTGGCAGCGGGGCGCGCTGCGCGGGGCTGCTTCGCA 839
Db 3808 CCGCTTCTTCGAGGAGTTCGTGCGCTGGCAGCGGGGCGCGCTGCGCGGGGCTGCTTCGCA 3867
QY 840 GTGGCGCGCTTTCACCGAGCGCGCGCTGCGCATGTTCCAGCGCTTACTGCGCCCTGGAGCC 899
Db 3868 GTGGCGCGCTTTCACCGAGCGCGCGCTGCGCATGTTCCAGCGCTTACTGCGCCCTGGAGCC 3927
QY 900 CGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 959
Db 3928 CGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3987
QY 960 GCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1019
Db 3988 GCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4047
QY 1020 CGGGCCACTTGGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGACGCGTCTGACCGAGAGCGG 1079
Db 4048 CGGGCCACTTGGCGCTCGAGCGCGCTGGCGCGCTCAAGCGGACGCGTCTGACCGAGAGCGG 4107
QY 1080 CAGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1139
Db 4108 CAGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4167
QY 1140 GGTGCCAGTCCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1199
Db 4168 GGTGCCAGTCCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4227
QY 1200 CGCGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259
Db 4228 CGCGCGGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4287
QY 1260 CGAGATCTCGTCTGA 1275
Db 4288 CGAGATCTCGTCTGA 4303

RESULT 10
US-09-734-032-3
; Sequence 3, Application US/09734032
; Patent No. US20020103116A1
; GENERAL INFORMATION:
; APPLICANT: WEI et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000536
; CURRENT APPLICATION NUMBER: US/09/734,032
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60205228
; PRIOR FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7301
; TYPE: DNA
; ORGANISM: human
US-09-734-032-3

Query Match 66.6%; Score 849.6; DB 10; Length 7301;
Best Local Similarity 99.5%; Pred. No. 4.9e-191;
Matches 852; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 420 CCCTCCCGAGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCT 479
Db 3448 CCGCCCGAGTGGGCTCCCTGAGGACACGCTGAAGCGCTGTGTCAGCAGCTGGGCT 3507
QY 480 GCGCTGGAATTCATGACGCGCGCGAGCTGTGTGACCGCGCATCAAGCCCGAGAACGT 539
Db 3508 GCGCTGGAATTCATGACGCGCGCGAGCTGTGTGACCGCGCATCAAGCCCGAGAACGT 3567
QY 540 GCTGCTGTGACCGCGAGTGCCTGTAAGCTGCGCGCATTCGCGCATGACGCGCG 599
Db 3568 GCTGCTGTGACCGCGAGTGCCTGTAAGCTGCGCGCATTCGCGCATGACGCGCG 3627
QY 600 GTCGGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 659
Db 3628 GTCGGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACACGCGCGCTGAGGTGTG 3687
QY 660 CCAGCGCGCGCGCGCGCGCTGGCGGTGACACGCGCGCTGTCAGCTGTGGGCTTCGG 719
Db 3688 CCAGCGCGCGCGCGCGCGCTGGCGGTGACACGCGCGCTGTCAGCTGTGGGCTTCGG 3747
QY 720 CGTGTCTATCTTCTCGTGTCTACCGGCAATCTTCCGTTGGAGGCGCGCTGGCGCGCG 779
Db 3748 CGTGTCTATCTTCTCGTGTCTACCGGCAATCTTCCGTTGGAGGCGCGCTGGCGCGCG 3807
```

Db	284	AGTTTGTGAACAGAGCAAAACCRAGCTTGAAGAACTTCTACGGGAGGTGAGCATCA	343
Qy	305	ACAGCCTCTCTCCAGCCCTTCATCATCAAGGCTTTGACGTGGTCTTTGAGACAGAGG	364
Db	344	ACAGCCTCTCTCCAGCCCTTCATCATCAAGGCTTTGACGTGGTCTTTGAGACAGAGG	403
Qy	365	ACTGCTAGCTTTTGCCCCAGAGTAGCACCCTGCTGGGGACCTGTTTGACATCATCCCTC	424
Db	404	ACTGCTAGCTTTTGCCCCAGAGTAGCACCCTGCTGGGGACCTGTTTGACATCATCCCTC	463
Qy	425	CCCAGGTGGGGCTCCCTCAGGACACGGTGAAGCGCTGTGCAGCAGCTGGCGCTGGCGC	484
Db	464	CCCAGGTGGGGCTCCCTCAGGACACGGTGAAGCGCTGTGTGACGAGCTGGGCCCTTGGCGC	523
Qy	485	TGGACTTCATGCACGGGCGGAGCTGTGTGCACCGCGACATCAAGCCCGAGAACGTGTGTC	544
Db	524	TGGACTTCATGCACGGGCGGAGCTGTGTGCACCGCGACATCAAGCCCGAGAACGTGTGTC	583
Qy	545	TGTTTGACCGGAGGTGCCCGCGCGTAAAGCTGTGCCGACTTCGGCATGACGCGCGCGTGG	604
Db	584	TGTTTGACCGGAGGTGCCCGCGCGTAAAGCTGTGCCGACTTCGGCATGACGCGCGCGTGG	643
Qy	605	GCTGCCGCGTCAAGCGGTGAGCGGCACATCTCCCTTACACGGCGCTCAGAGTGTGCCAGG	664
Db	644	GCTGCCGCGTCAAGCGGTGAGCGGCACATCTCCCTTACACGGCGCTCAGAGTGTGCCAGG	703
Qy	665	CGGCGCGCGCGACGGGCTGCGGTGGACACGGGCGTGGACGCTGTGGCGCTTCGCGCGTGC	724
Db	704	CGGCGCGCGCGACGGGCTGCGGTGGACACGGGCGTGGACGCTGTGGCGCTTCGCGCGTGC	763
Qy	725	TCATCTTCTGGGTGCTCACCGGCAACTTCCTCCGTGGGAGGCGCGTTCGGGCGCGAGCGCT	784
Db	764	TCATCTTCTGGGTGCTCACCGGCAACTTCCTCCGTGGGAGGCGCGTTCGGGCGCGAGCGCT	823
Qy	785	TCCTTCGAGGAGTTCGTGGCTGGCAGCGGGGCGCGCTGCCGGGGTGCTCTTCGAGTGGC	844
Db	824	TCCTTCGAGGAGTTCGTGGCTGGCAGCGGGGCGCGCTGCCGGGGTGCTCTTCGAGTGGC	883
Qy	845	GCOCCTTACCGAGCCCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTTGAGCCCGAGC	904
Db	884	GCOCCTTACCGAGCCCGCGCTGCGCATGTTCCAGCGCTTACTGGCCCTTGAGCCCGAGC	943
Qy	905	GCOCGGCCGACCAAGAGAGTGTTCGCTTCTCAAGACGAGTCACTGTCGAGTGTGC	964
Db	944	GCOCGGCCGACCAAGAGAGTGTTCGCTTCTCAAGACGAGTCACTGTCGAGTGTGC	1003
Qy	965	GCOCGGCCCTCGCACCGCGCGCAGCCCGCGGGACCGCCCGCCCGCCCGCGCGGC	1024
Db	1004	GCOCGGCCCTCGCACCGCGCGCAGCCCGCGGGACCGCCCGCCCGCCCGCGCGGC	1063
Qy	1025	CATTCGCGCTCGAGGCGCTTGGGCGCTCAAGCGGACGGTGTCTGACCGAGAGCGGACGC	1084
Db	1064	CATTCGCGCTCGAGGCGCTTGGGCGCTCAAGCGGACGGTGTCTGACCGAGAGCGGACGC	1123
Qy	1085	GCTCCCGGCGCGCGCCCGCTCGGCTCGGCTCGGCTTTCGCCGTGTCGCGTGTGCGGTGC	1144
Db	1124	GCTCCCGGCGCGCGCCCGCTCGGCTCGGCTCGGCTTTCGCCGTGTCGCGTGTGCGGTGC	1183
Qy	1145	CAGTGCCTGTGCGGTGCTGTGCCCGAGCCCGGCTAGCTCCCGAGGGGCCCCCGCGC	1204
Db	1184	CAGTGCCTGTGCGGTGCTGTGCCCGAGCCCGGCTAGCTCCCGAGGGGCCCCCGCGC	1243
Qy	1205	GGACCGACGGCCGCGGACAAAGGACAAAGGACAGGTGGTGTCTGGCCACGGCCATCGAGA	1264
Db	1244	GGACCGACGGCCGCGGACAAAGGACAAAGGACAGGTGGTGTCTGGCCACGGCCATCGAGA	1303
Qy	1265	TCTGGTCTGA	1275
Db	1304	TCTGGTCTGA	1314

US-09-916-790-3
; Sequence 3, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Slios-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,543
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-916-790-3

Query Match 81.6%; Score 1041; DB 9; Length 1041;
Best Local Similarity 100.0%; Pred. No. 2,7e-236;
Matches 1041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	235	ATGGCAGCTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCCTACGGAGGTG	294
DB	1	ATGGCAGCTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCCTACGGAGGTG	60
QY	295	AGCATCACCAAGCGCTCTCTCAGCCCTTTCATCATCAAGGTCCTTGAAGTGGTCTTT	354
DB	61	AGCATCACCAAGCGCTCTCTCAGCCCTTTCATCATCAAGGTCCTTGAAGTGGTCTTT	120
QY	355	GAGACGAGAGACTGCTACGTCCTTTCGCCAGAGTAGCGACCTGCTGGGACCTGTTTGAC	414
DB	121	GAGACGAGAGACTGCTACGTCCTTTCGCCAGAGTAGCGACCTGCTGGGACCTGTTTGAC	180
QY	415	ATCATCCCTCCCCAGGTGGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTCAGCAGCTG	474
DB	181	ATCATCCCTCCCCAGGTGGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTCAGCAGCTG	240
QY	475	GGCTGGCGCTGGACTTCATGCAAGCGGCGGACCTGGTGCAACCGGACATCAAGCCCGAG	534
DB	241	GGCTGGCGCTGGACTTCATGCAAGCGGCGGACCTGGTGCAACCGGACATCAAGCCCGAG	300
QY	535	AAGTGTGCTGTGTGACACCGGAGTGCAGCGCGGTAAAGCTGGCCGACTTCGGCGCTGACG	594
DB	301	AAGTGTGCTGTGTGACACCGGAGTGCAGCGCGGTAAAGCTGGCCGACTTCGGCGCTGACG	360
QY	595	CGCGCGTGGCTGCGCGCTGAAGCGCTGTAGCGGACCATCCCTTACACGGCGCCTGAG	654
DB	361	CGCGCGTGGCTGCGCGCTGAAGCGCTGTAGCGGACCATCCCTTACACGGCGCCTGAG	420
QY	655	GTGTGCACGCGCGCGCGCGCGCTGGCGGTGGACACGGGCGGTGACGCTGTGGGCC	714
DB	421	GTGTGCACGCGCGCGCGCGCGCTGGCGGTGGACACGGGCGGTGACGCTGTGGGCC	480
QY	715	TTGGGCGTGTCTATCTTCTGCTGTCTACACGGGCAACTTCCCGTGGAGAGCGCGCTGGGC	774
DB	481	TTGGGCGTGTCTATCTTCTGCTGTCTACACGGGCAACTTCCCGTGGAGAGCGCGCTGGGC	540
QY	775	GCCGAGCGCTTCTTCAGAGAGTTCTGGTGGCTGCGACGGGGCGCGCTGCCGGGGCTGCCT	834
DB	541	GCCGAGCGCTTCTTCAGAGAGTTCTGGTGGCTGCGACGGGGCGCGCTGCCGGGGCTGCCT	600
QY	835	TGCGAGTGGCGCGCTTACCGAGCGCCGGCTGCGCATGTTCCAGCGCTTACTTGGCCCTG	894
DB	601	TGCGAGTGGCGCGCTTACCGAGCGCCGGCTGCGCATGTTCCAGCGCTTACTTGGCCCTG	660
QY	895	GAGCCGAGCGCGCGCGCGCGCGCCAGCCAAAGTAGTGTTCGGTTCCTCAAGCACGAGCTCACG	954
DB	661	GAGCCGAGCGCGCGCGCGCGCGCCAGCCAAAGTAGTGTTCGGTTCCTCAAGCACGAGCTCACG	720

OTHER INFORMATION: Incyte ID No. US20030207299A1 7312543CB1
US-10-288-798-25

Query Match 99.9%; Score 1273.4; DB 12; Length 2060;
Best Local Similarity 99.9%; Pred. No. 4.1e-291;
Matches 1274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGCGTGGGCTCCAGAGCCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGCCCGGGG 60
DB 445 ATGAGCGTGGGCTCCAGAGCCCTGAGCGCGCCCGCTCCCTGACCTGCTGTGGCCCGGGG 504
QY 61 ACTCCCTCCCTGGGCTGTGGCGGTGTGCCCTTCTCACTGAAGACATGACGAGCCCTGACT 120
DB 505 ACTCCCTCCCTGGGCTGTGGCGGTGTGCCCTTCTCACTGAAGACATGACGAGCCCTGACT 564
QY 121 CTCGGCACACTGGCGCGCCAGCGAGCTGACCAAGCACTAGAACCTAGTTCGGGAGCTGGG 180
DB 565 CTCGGCACACTGGCGCGCCAGCGAGCTGACCAAGCACTAGAACCTAGTTCGGGAGCTGGG 624
QY 181 AAAGGCACCTATGGGAAGGTTGACCTGGTGGTCTACAAGGCGACAGGCACAAAATGGCA 240
DB 625 AAAGGCACCTATGGGAAGGTTGACCTGGTGGTCTACAAGGCGACAGGCACAAAATGGCA 684
QY 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 300
DB 685 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACGGAGGTGAGCATC 744
QY 301 ACCAACAGCCTCTCTCCAGCCCTTCAATCATCAAGCTTTTGAAGTCTTTTGAACA 360
DB 745 ACCAACAGCCTCTCTCCAGCCCTTCAATCATCAAGCTTTTGAAGTCTTTTGAACA 804
QY 361 GAGGACTGTACGTTCTTGGCCAGAGTACGACCTGCTGGGACCTGTTTGAATCATC 420
DB 805 GAGGACTGTACGTTCTTGGCCAGAGTACGACCTGCTGGGACCTGTTTGAATCATC 864
QY 421 CCTCCCAAGTGGGCTCCCTGAGGACACAGCTGAAGGCTGTGTGACGACCTGGGCTG 480
DB 865 CCTCCCAAGTGGGCTCCCTGAGGACACAGCTGAAGGCTGTGTGACGACCTGGGCTG 924
QY 481 CGCTGGAATTCATGACAGGCGGAGCTGTGTGACCGGACATCAAGCCGAGAACTGTG 540
DB 925 CGCTGGAATTCATGACAGGCGGAGCTGTGTGACCGGACATCAAGCCGAGAACTGTG 984
QY 541 CTGCTGTTCGACCGGAGTCCGCGGTGAAGCTGGCCGACTTCGCATGACCGCGGC 600
DB 985 CTGCTGTTCGACCGGAGTCCGCGGTGAAGCTGGCCGACTTCGCATGACCGCGGC 1044
QY 601 GTGGGCTGCCGCGTCAAGCGCGTGAAGGCGCACCATCCCTTACAGCGGCGCTGAGGTG 660
DB 1045 GTGGGCTGCCGCGTCAAGCGCGTGAAGGCGCACCATCCCTTACAGCGGCGCTGAGGTG 1104
QY 661 CAGGCGGCGCGCGGAGTGGCGGTGACACAGGCGCTGGAGCTGTGGGCCCTTGGC 720
DB 1105 CAGGCGGCGCGCGGAGTGGCGGTGACACAGGCGCTGGAGCTGTGGGCCCTTGGC 1164
QY 721 GTGCTCATCTTCTGCTGCTCACCGGCACTTCCGTGGGAGGCGGCTGGGCGCGGAC 780
DB 1165 GTGCTCATCTTCTGCTGCTCACCGGCACTTCCGTGGGAGGCGGCTGGGCGCGGAC 1224
QY 781 GCCTTCTTCGAGGAGTTCGTGCTGGCAGCGGGGCGGCTTCCGCGGCGTCCCTTCGAG 840
DB 1225 GCCTTCTTCGAGGAGTTCGTGCTGGCAGCGGGGCGGCTTCCGCGGCGTCCCTTCGAG 1284
QY 841 TGGCGCGCTTCACGAGCGCGCTGCGATATTTTCAGGCGCTTACTGGGCCCTGGAGCC 900
DB 1285 TGGCGCGCTTCACGAGCGCGCTGCGATATTTTCAGGCGCTTACTGGGCCCTGGAGCC 1344
QY 901 GAGCGCGCGCGCGAGCAAGAGGTGTTCCGCTTCTCTCAAGCAGAGCTCACCTCCGAG 960
DB 1345 GAGCGCGCGCGCGAGCAAGAGGTGTTCCGCTTCTCTCAAGCAGAGCTCACCTCCGAG 1404
QY 961 CTGCGCGCGCGCGCTTCGACACCGCGCGGCAAGGCGGCGGCGCGCGCGCGCGCG 1020
```

```
DB 1405 CTGCGCGCGCGCGCTGCGACCCGCGCGCGCAAGCCCGCGGGAACCGCCCGCGCGCGC 1464
QY 1021 GGGCCACTGCGCTCGAGCGCGCTGGGCGCGCTCAAGCGGACGGTGTGTGACCGGAGCGGC 1080
DB 1465 GGGCCACTGCGCTCGAGCGCGCTGGGCGCGCTCAAGCGGACGGTGTGTGACCGGAGCGGC 1524
QY 1081 AGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 1525 GCGCGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1584
QY 1141 GTCCAGTGCCTGCGCGGTGCTGTGCGGAGCCCGCGCGCGCGCGCGCGCGCGCGCG 1200
DB 1585 GTCCAGTGCCTGCGCGGTGCTGTGCGGAGCCCGCGCGCGCGCGCGCGCGCGCGCG 1644
QY 1201 GCGCGGACCGACCGCGCGCGCGGACAAAGGAGGAGGTGTGTGCGGACCGGCGCATC 1260
DB 1645 GCGCGGACCGACCGCGCGCGCGGACAAAGGAGGAGGTGTGTGCGGACCGGCGCATC 1704
QY 1261 GAGATCTGCGTCTGA 1275
DB 1705 GAGATCTGCGTCTGA 1719
```

RESULT 7
US-09-916-790-1
; Sequence 1, Application US/09916790
; Patent No. US20020061573A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachael
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 381552002700
; CURRENT APPLICATION NUMBER: US/09/916,790
; PRIOR FILING DATE: 2001-07-27
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2893
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (274)...(1314)
; NAME/KEY: misc feature
; LOCATION: (1)...(2893)
; OTHER INFORMATION: n = A, T, C or G
US-09-916-790-1

Query Match 88.8%; Score 1131.8; DB 9; Length 2893;
Best Local Similarity 94.8%; Pred. No. 1.1e-257;
Matches 1205; Conservative 0; Mismatches 57; Indels 9; Gaps 3;

```
QY 14 GCCCAGAGCCTGAGCGCGCGCGCTCCCTGACCTGTGTGGGCGCGGAGCTGCCCTTGGGC 73
DB 44 GCCTTGAACCGCGCGCGCTCCCTGACCTTGTGTGGCGCGGAGGACTGCCCTCTGGGC 103
QY 74 CTGGTGGCG--GTGTGCGCTTCTCACTGAAGACATGACGCGCGCTGACTCTCCGACAC 130
DB 104 CTGGTCCGGGTGTGCCCTTCTCACTTGAAGACATGACGCGCGCTGACTCTCCGACAC 163
QY 131 TGG--CCGCGCAGCGAGTCAACCAAGCACTACGAACTAGTCCCGGAGCTGGGCAAG 184
DB 164 TTGGCCCGCGCAGCGAGTCAACCAAGCACTACGAACTAGTCCCGGAGCTGGGCAAG 223
QY 185 GCACCTATGGGAAGTGTGACCTGGTGTCTAAGGGGACAGGCACAAAATGGCACTGA 244
DB 224 GCACCTACGGGAAGTGTGATCTGGTGTCTAAGGGGACAGGCACAAAATGGCACTGA 283
QY 245 AGTTTGTGAACAGAGGCAAAACCAAGCTGAAGAACTTCCTACGGGAGGTGAGCATCACA 304
```


Qy	181	AAAGGCACCTATGGGAAAGGTTGACCTGGTGGTCTTA	CAAAGGGCA	CAGGCA	CAAAAAATGGCA	240
Db	181	AAAGGCACCTATGGGAAAGGTTGACCTGGTGGTCTTA	CAAAGGGCA	CAGGCA	CAAAAAATGGCA	240
Qy	241	CTGAAGTTTGTGAACACAGAGCAAAACCAAGCTGA	AGAACTCTCTACGGAGGTGAGCATC	300		
Db	241	CTGAAGTTTGTGAACACAGAGCAAAACCAAGCTGA	AGAACTCTCTACGGAGGTGAGCATC	300		
Qy	301	ACCAACAGCCTCTCCTCCAGCCCTTCATCATCAAG	GTCTTTGACGATCATCATC	360		
Db	301	ACCAACAGCCTCTCCTCCAGCCCTTCATCATCAAG	GTCTTTGACGATCATCATC	360		
Qy	361	GAGGACTGTACGTCCTTTGCCAGGAGTACGACCT	GCTGGGAGCATCTTTTGACATCATC	420		
Db	361	GAGGACTGTACGTCCTTTGCCAGGAGTACGACCT	GCTGGGAGCATCTTTTGACATCATC	420		
Qy	421	CCTCCCAGGTGGGGCTCCCTTGAGGACACGGTGA	AGCGCTGTGTGACGACAGCTTGGGGCTG	480		
Db	421	CCTCCCAGGTGGGGCTCCCTTGAGGACACGGTGA	AGCGCTGTGTGACGACAGCTTGGGGCTG	480		
Qy	481	GGCTGGACTTCATGCA	CGGCGGCAGCTGGTGACCGGACATCAAG	540		
Db	481	GGCTGGACTTCATGCA	CGGCGGCAGCTGGTGACCGGACATCAAG	540		
Qy	541	CTGCTGTTCGACCGCGAGTGC	CGCGCGGTAAAGCTGCGCGACTTCGCGCATGA	600		
Db	541	CTGCTGTTCGACCGCGAGTGC	CGCGCGGTAAAGCTGCGCGACTTCGCGCATGA	600		
Qy	601	GTGGGCTGCCGCGTCAAGCGCGT	GAGCGGACCATCCCTTACACGGCGCTGAGGTGTC	660		
Db	601	GTGGGCTGCCGCGTCAAGCGCGT	GAGCGGACCATCCCTTACACGGCGCTGAGGTGTC	660		
Qy	661	CAGCGGGCGCGCGCGAGCTGGCGGTGGACACG	GCGGTGGACGTGTGGGCTTCGCG	720		
Db	661	CAGCGGGCGCGCGCGAGCTGGCGGTGGACACG	GCGGTGGACGTGTGGGCTTCGCG	720		
Qy	721	GTGCTCATCTTCTCGGTGCTC	ACCGGCAACTCTCCGCTGGAGGCGGCGT	780		
Db	721	GTGCTCATCTTCTCGGTGCTC	ACCGGCAACTCTCCGCTGGAGGCGGCGT	780		
Qy	781	GCCTTCTTCGAGGAGTTCGTGCGCTGGCAGCG	GGGCGCGCTTCGCTTCGCAG	840		
Db	781	GCCTTCTTCGAGGAGTTCGTGCGCTGGCAGCG	GGGCGCGCTTCGCTTCGCAG	840		
Qy	841	TGGCGCGCTTTCACGAGCCGCGTGGCGATGTTC	AGGCGCTTACCTGGAGGCC	900		
Db	841	TGGCGCGCTTTCACGAGCCGCGTGGCGATGTTC	AGGCGCTTACCTGGAGGCC	900		
Qy	901	GAGCGCGCGCGCCAGCCAAAGAGGTGTTCGCT	TTCCTCAGCAGCAGCTACGCTCCGAG	960		
Db	901	GAGCGCGCGCGCCAGCCAAAGAGGTGTTCGCT	TTCCTCAGCAGCAGCTACGCTCCGAG	960		
Qy	961	CTGGCGCGCGCGCTTCGACACCGCGCGCAAGC	CCCCCGGGAGCCCGCGCGCGCGCC	1020		
Db	961	CTGGCGCGCGCGCTTCGACACCGCGCGCAAGC	CCCCCGGGAGCCCGCGCGCGCGCC	1020		
Qy	1021	GGGCGACTGTGCGCTCGAGGCGCTTGGGGCGCT	CAAGCGGACGCTGTGACGAGAGCGGC	1080		
Db	1021	GGGCGACTGTGCGCTCGAGGCGCTTGGGGCGCT	CAAGCGGACGCTGTGACGAGAGCGGC	1080		
Qy	1081	AGCGCTCCCGGCGCGCGCCCCCGCGCTCGGGT	CGGTGCGCTTGGCGCTGCGCGCGGTCCG	1140		
Db	1081	AGCGCTCCCGGCGCGCGCCCCCGCGCTCGGGT	CGGTGCGCTTGGCGCTGCGCGCGGTCCG	1140		
Qy	1141	GTGCCAGTGTCCCGTGCCTGTGCGCGAGCCCG	CGCTTAGCTCCCCAGGGGCGCCCC	1200		
Db	1141	GTGCCAGTGTCCCGTGCCTGTGCGCGAGCCCG	CGCTTAGCTCCCCAGGGGCGCCCC	1200		
Qy	1201	GGCCGGAACCGACGGCGCGCGGACAAAGAGCA	AAAGGGCAGGTGGTGTGGCA	1260		
Db	1201	GGCCGGAACCGACGGCGCGCGGACAAAGAGCA	AAAGGGCAGGTGGTGTGGCA	1260		

```

QY      1261 GAGATCTGGCTCTGA 1275
      |||||
Db      1261 GAGATCTGGCTCTGA 1275

RESULT 2
US-10-016-985-3
; Sequence 3, Application US/10016985
; Publication No. US20020123621A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Maricar, Miranda
; APPLICANT: Yu, Xuanchuan (Sean)
; APPLICANT: Fiddle, Carl Johan
; TITLE OF INVENTION: No. US20020123621A1el Human Kinase and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-02773-USA
; CURRENT APPLICATION NUMBER: US/10/016,985
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/251,941
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-985-3

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 16:12:04 / Search time 334 Seconds
(without alignments)
12687.392 Million cell updates/sec

Title: US-10-016-985-1
Perfect score: 1275
Sequence: 1 atgagctgggtcccgaga.....ccatcgagatctggtctga 1275

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

```
1: /cgn2_6/ptodata/2/pubnpa/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubnpa/ECT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubnpa/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubnpa/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubnpa/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubnpa/ECTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubnpa/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubnpa/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubnpa/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubnpa/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubnpa/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubnpa/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubnpa/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubnpa/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubnpa/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubnpa/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubnpa/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1275	100.0	1275	13	US-10-016-985-1
2	1275	100.0	1473	13	US-10-016-985-3
3	1275	100.0	2598	9	US-09-816-094-1
4	1275	100.0	2598	14	US-10-233-613-1
5	1273.4	99.9	1275	10	US-09-734-032-1
6	1273.4	99.9	2060	12	US-10-288-798-25
7	1131.8	88.8	2893	9	US-09-916-790-1
8	1041	81.6	1041	9	US-09-916-790-3
9	849.6	66.6	7301	9	US-09-816-094-3
10	849.6	66.6	7301	14	US-09-734-032-3
11	849.6	66.6	7301	14	US-10-233-613-1
12	767.6	60.2	788	9	US-09-861-893-38
13	236.6	18.6	1247	12	US-10-353-690-49
14	235	18.4	1156	12	US-10-288-798-40
15	99.8	7.8	1257	9	US-09-799-875-15
16	99.8	7.8	1260	12	US-10-303-664A-5

```

17 99.8 7.8 1826 9 US-09-799-875-13 Sequence 13, Appli
18 99.8 7.8 1826 12 US-10-303-664A-4 Sequence 4, Appli
19 99.8 7.8 1864 11 US-09-829-607-1 Sequence 1, Appli
20 99.2 7.8 2978 11 US-09-955-999-43 Sequence 43, Appli
21 98.4 7.7 1854 14 US-10-198-846-9984 Sequence 9984, Ap
22 98.2 7.7 1972 13 US-10-098-841-166 Sequence 166, App
23 97.6 7.7 2181 14 US-10-072-036-70 Sequence 70, Appli
24 97.6 7.7 2184 14 US-10-072-036-138 Sequence 138, App
25 97.6 7.7 2610 10 US-09-970-000-3 Sequence 3, Appli
26 93.4 7.3 1599 12 US-10-324-985A-3 Sequence 254, App
27 93.4 7.3 1715 12 US-10-007-926A-254 Sequence 5, Appli
28 89.4 7.0 1140 12 US-10-324-985A-5 Sequence 1, Appli
29 89.4 7.0 2626 12 US-10-324-985A-1 Sequence 1, Appli
30 88.8 7.0 1635 10 US-09-880-107-2340 Sequence 2340, Ap
31 87.2 6.8 9025608 14 US-10-156-761-6069 Sequence 6069, Ap
32 87.2 6.8 9025608 14 US-10-156-761-1 Sequence 1, Appli
33 86.2 6.8 819 14 US-10-029-386-25154 Sequence 25154, A
34 86.2 6.8 858 12 US-10-029-386-25154 Sequence 1, Appli
35 86.2 6.8 1149 12 US-09-863-776-1 Sequence 1, Appli
36 86.2 6.8 1376 10 US-09-764-868-237 Sequence 237, App
37 86.2 6.8 32192 10 US-09-764-847-1416 Sequence 1416, Ap
38 86.2 6.8 32192 11 US-09-764-891-7945 Sequence 7945, Ap
39 86.2 6.8 32192 11 US-09-764-891-8220 Sequence 8220, Ap
40 86.2 6.8 32192 11 US-09-764-891-8319 Sequence 8319, Ap
41 86.2 6.8 32192 14 US-10-092-154-1416 Sequence 1416, Ap
42 85.8 6.7 1155 14 US-10-156-761-4992 Sequence 4992, Ap
43 85.8 6.7 9025608 14 US-10-156-761-1 Sequence 1, Appli
44 85.4 6.7 3270 10 US-09-835-081-1 Sequence 1, Appli
45 85.2 6.7 1368 10 US-09-764-868-557 Sequence 557, App

```

ALIGNMENTS

RESULT 1

```

; Sequence 1, Application US/10016985
; Publication No. US20020123621A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Maricar, Miranda
; APPLICANT: Yu, Xuanchuan (Sean)
; APPLICANT: Fridele, Carl Johan
; TITLE OF INVENTION: No. US20020123621A1 Human Kinase and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0273-USA
; CURRENT APPLICATION NUMBER: US/10/016,985
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/251,941
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1275
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-016-985-1

```

```

Query Match 100.0%; Score 1275; DB 13; Length 1275;
Best Local Similarity 100.0%; Pred. No. 1.6e-291;
Matches 1275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGTGGCTGCCAGAGCCTGAGCGCGCGCTCCCTGACCTCTGTGGGCGCGGG 60
Db 1 ATGAGCGTGGCTGCCAGAGCCTGAGCGCGCGCTCCCTGACCTCTGTGGGCGCGGG 60

QY 61 ACTCCCTGGCTGTGGCGGTGTCCTTCTACTGAAGACATGCGAGCCCTGACT 120
Db 61 ACTCCCTGGCTGTGGCGGTGTCCTTCTACTGAAGACATGCGAGCCCTGACT 120

QY 121 CTCGCGACACTGGCGCGCGAGCGAGTACCAAGCACTACGAACCTAGTCCGAGCTGGC 180
Db 121 CTCGCGACACTGGCGCGCGAGCGAGTACCAAGCACTACGAACCTAGTCCGAGCTGGC 180

```

gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGCTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 135 a 206 c 260 g 165 t 1 others
ORIGIN

Query Match 46.6%; Score 594.4; DB 13; Length 767;
Best Local Similarity 88.8%; Pred. No. 3.3e-105;
Matches 677; Conservative 0; Mismatches 81; Indels 4; Gaps 3;

QY 124 CGCACACTGGCGCGCAGCGCTACCAAGCACTACGAACTAGTCCGGGAGCTGGGCAAA 183
DB 1 CGCACACTGGCTGCCAGCGATGTTACCAAGCACTATGAGCTTGTGGGAGCTGGGCAAA 60

QY 184 GGCACCTATGGGAAGGTTGACCTGTGTGTTCTACAAGGCGACAGGCACAAAATGGCACTG 243
DB 61 GGGACCTACGGGAAGGTCGACCTGTGTGCTTACAAAGGCGACAGGCACCTAAAATGGCCCTG 120

QY 244 AAGTTGTGAACAGAGCAAAACCAAGCTGAGAACCTTCTACGGGAGGTGAGCATCACC 303
DB 121 AAAATTGTGAATAAGAGTAAGACAAAGCTGAAGAACTTCTGCGGTGAGGTGAGCATCACC 180

QY 304 AACAGCCCTCTCCTCCAGGCCCTTTCATCATCAAGGTCTTTGACGTGTCTTTGAGACAGAG 363
DB 181 AACAGCCCTGTGCTGTAGCCCTTTCATCATCAAGGTCTTTGACGTGTCTTTGAGACCCAG 240

QY 364 GACTGCTACGCTTTTGGCAGAGTACGCACTGCTGGGACCTGTTTGACATCATCCCT 423
DB 241 GAGTGTCTATGTCTTTGGCCAGAGTATGACCTGCTGGGACCTGTTTGACATCATCCCT 300

QY 424 CCCAGGTGGGGCTCCCTGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGCCCTGGCG 483
DB 301 CCTCAGGTGGGGCTCCCGAGAGGACACGGTGAAGCGCTGTGTGACAGCTGGGGCTGGCG 360

QY 484 CTGGACTTCATGCACGGCGGCGAGCTGTGTGACCCGACATCAAGCCCGAGAACGTGCTG 543
DB 361 CTGGACTTCATGCACAGCAGGAGCTGTGTGACCCGACATCAAGCCCGAGAACGTGCTG 420

QY 544 CTGTTGACCGCGAGTGCAGCGCGTAAAGTGGCCGACTTGGGATGACGCGCGCGTG 603
DB 421 CTGTTGACCGTGAAGTGCAGCGCGTGAAGCTGGCTGACTTCGGCATGACGCGCGCGTG 480

QY 604 GGCTGCGCGTCAAGCGGTGAGCGGACCATCCCTTACAGGCGCTGAGGTGTGCCAG 663
DB 481 GGCTGCGCGTGAAGCGGTGAGCGGCACTATACCGTACACAGCACCCGAGGTGTGCCAG 540

QY 664 GCGGCGCGCGCGAGCGGCTGGCGGTGACACGGGCGTGGACGTGTGGGCCCTTCGGCGTG 723
DB 541 GCTGCGCGCGCGATGGCTTCGGGTGACACGGGCGTGGACGTGTGGGCCCTTTGGCGTG 600

QY 724 CTCATCTTCTCGGTGCTCACCGGCAACTTCC-CGTGGAGGCGGCGTGGGCGCGGACGC 782
DB 601 CTCATCTTCTGTGTGCTCACTGGCAACTTCNCGTGGGAGGCTGATCGGCTGGGATGC 660

QY 783 CTTCTTCAGGAGTTCGTGCGCTGCGAGCGGGCGCTGCGGGGCTGCCCTTCGCACTG 842
DB 661 CTTCTTCAGGAGTTCGTGCGCTGCGAGCGGGGCTGCCCTTCGCGGGCTGCC-ACGCACTG 719

QY 843 GCGCGCTTTCACCGAGCGCGGCTGCGCATGTTCAGCGGCTT 884
DB 720 GCGC--GTTTACCGAGCTGCGCTGCGTATGTTCAGCGGCTT 759

Search completed: December 7, 2003, 16:46:04
Job time : 2030 secs

BQ179888.1 GI:20355380
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 706)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5703871"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DHI0B (T1 phage resistant)"
 /clone_lib="NIH-BMAP_EW0"
 /note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GTCGCTGGGA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."
 BASE COUNT 109 a 203 c 242 g 152 t
 ORIGIN
 Query Match 46.8%; Score 596.8; DB 13; Length 706;
 Best Local Similarity 90.5%; Pred. No. 1.1e-105;
 Matches 637; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 QY 265 ACCAAGCTGAAGACTTCTCAGCGAGGTGAGCATCCACACAGCCTCTCTCCAGCCCC 324
 DB 2 ACAAGCTGAAGACTTCTCAGCGAGGTGAGCATCCACACAGCCTCTCTAGCCCC 61
 QY 325 TTCATCATCAAGGCTTTTACGCTGCTTTGAGACAGAGGACTGCTAGCTCTTTGCCCG 384
 DB 62 TTCATCATCAAGGCTTTTACGCTGCTTTGAGACCGAGGAGTCTATGCTTTGCCCG 121
 QY 395 GAGTACGACCTGCTGGGAGCCTGTTTGACATCATCTCCCTCCCGAGTGGGCTCCCTGAG 444
 DB 122 GAGTATGACCTGCTGGGAGCCTGTTTGACATCATCTCCCTCCCGAGTGGGCTCCCGAG 181
 QY 445 GACAGGTGAAGCGCTGTGCAGCAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 504
 DB 182 GACAGGTGAAGCGCTGTGTGAGCAGCAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 241
 QY 505 CAGCTGTGTGACCGCGCATCAAGCCCGAGAACGCTGCTGCTGTTTCGACCGGAGTGGCGC 564

Db 242 CAGCTGTGTGACCGTGACATCAAGCCCGAGATGCTGCTGTGTTGACCGTGTGCGC 301
 QY 565 CGGCTAAAGCTGGCCGACTTTCGGCATGACCGCGCGCTGGGCTGCCCGCTCAAGCGGTG 624
 Db 302 CGTGTGAAGCTGGCTGACTTCGGCATGACGGGGCGCTGGGCTGCCGTGTGAAGCGGTG 361
 QY 625 AGCGGACCACTCCCTTACAGCGGCGCTGAGTGTGCCAGGGGCGCGCGGAGCGGTG 684
 Db 362 AGCGGACCACTATACCGTACACAGCACCGCGAGTGTGCCAGGCTGGCGCGCGGATG 421
 QY 685 GCGGTGACACGCGCGGTGGACGTGTGGCGCTTGGCGGTGCTCATCTTCTGCTGCTCAC 744
 Db 422 GCGGTGACACGCGCGGTGGACGTGTGGCGCTTGGCGGTGCTCATCTTCTGCTGCTCAC 481
 QY 745 GGCAACTTCCGCTGGGAGGCGGTGCGGCGCGCGACGCCCTTTCGAGGAGTGTGCTGCGC 804
 Db 482 GGCAACTTCCGCTGGGAGGCTGTCATCGGTGCGGATGCGATGCTTTCGAGGAGTGTGCTGCGC 541
 QY 805 TGGCAGCGGGCGCGCTTCCGCGGCGCTTCCGAGTGGCGCGCTTTCACGAGCGCGCG 864
 Db 542 TGGCAGCGGGCGCGCTTCCGCGGCGCTTCCGAGTGGCGCGCTTTCACGAGCGCGCG 601
 QY 865 CTGCGCATGTTCCAGCGCTTACTGGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGAG 924
 Db 602 CTGCGTATGTTCCAGCGGCTTCTGGCGCTAGAGCTGAGCGCGCGCGCGCGCGCGCGAG 661
 QY 925 GTGTTCCGCTTCCCTCAAGCAGAGCTCACTGCTGCGAGTGGCGCGCGCGCGCGCGCG 968
 Db 662 GTCTTTCGCTTCCCTCAAGCAGAGCTCACTGCTGCGAGTGGCGCGCGCGCGCGCGCG 705
 RESULT 15
 BU613351 767 bp mRNA linear EST 20-FEB-2003
 LOCUS UI-M-EW0-cay-p-20-0-UI.r1 NIH-BMAP_EW0 Mus musculus cDNA clone
 DEFINITION UI-M-EW0-cay-p-20-0-UI 5', mRNA sequence.
 ACCESSION BU613351
 VERSION BU613351.1 GI:23279566
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 767)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bent-soares@uiowa.edu
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..767
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="UI-M-EW0-cay-p-20-0-UI"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DHI0B (T1 phage resistant)"
 /clone_lib="NIH-BMAP_EW0"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose

FEATURES
source

QY 508 CTGTGTCACCGCGACATCAAGCCGAGAACTGTGTGTCGTCGACGCGAGTGCCTGCGCG 567
 Db 421 CTGTGTCACCGCGACATCAAGCCGAGAACTGTGTGTCGTCGACGCGAGTGCCTGCGCG 480
 QY 568 GTAAAGCTGGCGACTTCGGCATCAGCGCGCGTGGCTGGCGTCCGCTCAGCGGCTGAGC 627
 Db 481 GTGAAGCTGGCTGACTTCGGCATCAGCGCGCGTGGCTGGCGTCCGCTGGAAGCGGTGAGC 540
 QY 628 GGCACCATCCCTTACACGGCGCCTGAGGTGTGCCAGCGCGCGCGCGCGCGCTGGCG 687
 Db 541 GGCATATACCGTACACAGCACCGAGGTGTGCCAGCGCGCGCGCGCGCGCTGGCG 600
 QY 688 GTGACACGGCGGTGGAGTGTGGCGCTTGGCGTGTCTATCTTCTGCTGTCTCAGCGCG 747
 Db 601 GTGACACGGCGGTGGAGTGTGGCGCTTGGCGTGTCTATCTTCTGCTGTCTCAGTGGC 660
 QY 748 AACTTCCGCTGGAGGGCGGTTCGGCGCGCGACGCTTCTTCGAGGAGTTCGTGGCTGG 807
 Db 661 AACTTCCGCTGGAGGGTGTGATCGGGTGGCGATG-CITCTTCGAGGAGTTCGTGGCTGG 719
 QY 808 CAGCGGGCGCG 818
 Db 720 CAGCGGGTGC 730

RESULT 13
 BM947152 733 bp mRNA linear EST 14-MAR-2002
 LOCUS UI-M-EH0P-bva-d-11-0-UI.r1 NIH_BMAP_EH0P Mus musculus cDNA clone
 DEFINITION IMAGE:5689306 5', mRNA sequence.
 ACCSSION BM947152
 VERSION BM947152.1 GI:19430737
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 733)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 source
 1. 733
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5689306"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EH0P"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then

cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 , is CAGCCACGAC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institute of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 113 a 216 c 249 g 154 t 1 others
 ORIGIN

Query Match 47.3%; Score 603.2; DB 12; Length 733;
 Best Local Similarity 89.8%; Pred. No. 6.4e-107;
 Matches 658; Conservative 0; Mismatches 74; Indels 1; Gaps 1;
 QY 265 ACCAAGCTGAAGAACTTCCTACGCGAGTGTGAGCATCAACACAGCTCTCTCTCAGCCCC 324
 Db 2 ACAAGCTGAAGAACTTCCTGCGTGAAGTGTGAGCATCAACACAGCTCTCTCTAGCCCC 61
 QY 325 TTCTATCATCAAGTCTTTGAGGTGGTCTTTGAGACAGAGGACTGCTACGCTTTTGCCAG 384
 Db 62 TTCTATCATCAAGTCTTTGAGGTGGTCTTTGAGACAGAGGAGTGTCTATGCTTTTGCCAG 121
 QY 385 GAGTACGACACCTGCTGGGACCTGTTTGACATCATCTCCCTCCACAGGTGGGCTCCCTGAG 444
 Db 122 GAGTATGACCTGCTGGGACCTGTTTGACATCATCTCCCTCTCAGGTGGGCTCCACAG 181
 QY 445 GACACGTGAAGCGCTGTGTGACAGACCTGGGCTGTGGCTGTGACTTTCATGCAAGCGCG 504
 Db 182 GACACGTGAAGCGCTGTGTGACAGACCTGGGCTGTGGCTGTGACTTTCATGCAAGCAGG 241
 QY 505 CAGCTGCTGACCGCGACATCAAGCCGAGACGTCGTGTGTCGACCGGCGCGCGAGTGGCG 564
 Db 242 CAGCTGCTGACCGGTGACATCAAGCCGCGAATGTGTGTGTTGACCGGTGAGTGGCG 301
 QY 565 CGCGTAAAGCTGGCGGACTTCGGCATCAGCGCGCGCTGGGCTGCCGCTCAAGCGCGTG 624
 Db 302 CGTGTGAAGTGGTGTGACTTCGGCATCAGCGCGCGCTGGGCTGCCGCTGTGAAGCGCGTG 361
 QY 625 AGCGGCAACCATTCCTTACAGCGCGCTGAGGTGTGCCAGCGCGCGCGCGCGAGTGGCG 684
 Db 362 AGCGGCACTATACCGTACAGCACCGCGGTGTGCCAGCGCGCGCGCGCGATGGCTTC 421
 QY 685 CGCGTGAACAGCGCGGTGAGGTGTGGGCTTGGGCTGTGCTCATCTTCTGGGCTCTAC 744
 Db 422 CGCGTGAACAGCGCGGTGAGCGTGTGGGCTTGGGCTGTGCTCATCTTCTGTGCTCACT 481
 QY 745 GGCACATTCCTCGTGGGAGCGCGCTGCCGCGCGCGCGCGCTTCTTCGAGGAGTTCGTGCG 804
 Db 482 GGCACATTCCTCGTGGGAGCGTGCATCGGTGCGATGCCCTTCTTCGAGGAGTTCGTGCG 541
 QY 805 TGGCAGCGGGCGCGCTGCGCGGCGCTGCTTTCGAGTGGCGCGCGCTTTCACGAGCGCGCG 864
 Db 542 TGGCAGCGGGCTGCGCTGCGCGGCGCTGCATCGCAGTGGCGCGCGCTTTTACGAGCGCTGC 601
 QY 865 CTGCGCATGTTCCAGCGCTTACTGCGCTTGAGCGCGCGCGCGCGCGCGCGCGCAAGGAG 924
 Db 602 CTGCGTATGTTCCAGCGCTTCTGCGCTTAGAGCGCGCGCGCGCGCGCGCGCGCAAGGAG 661
 QY 925 GTGTTCCGCTTCTCCTCAAGCAGCGCTCAGCGTCCGAGTGGCGCGCGCGCGCGCTTCGACCGCG 984
 Db 662 GTCTTTGCTTCTCCTCAAGCAGCGCTCAGCATCCGAGTGGCGCGCGCGCGCTTCNCAC-CGC 720
 QY 985 GCGCGCAAGCGCG 997
 Db 721 GCAGCAAGCGCG 733

RESULT 14
 BQ179888 706 bp mRNA linear EST 30-APR-2002
 LOCUS UI-M-EW0-bww-c-08-0-UI.r1 NIH_BMAP_EW0 Mus musculus cDNA clone
 DEFINITION IMAGE:5703871 5', mRNA sequence.
 ACCESSION BQ179888

RESULT 12	BU701617	730 bp	linear	EST 09-OCT-2002
LOCUS	BU701617			
DEFINITION	UI-M-FIO-byp-i-05-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone			
	IMAGE: 5721292 5', mRNA sequence.			
ACCESSION	BU701617			
VERSION	BU701617.1	GI:23625601		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 730)			
AUTHORS	NTH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			

BASE COUNT	139	a	197	c	241	g	153	t	ORIGIN	
Query Match	47.4%	Score	603.8	DB	13	Length	730			
Best Local Similarity	90.0%	Pred. No.	4.9e-107							
Matches	658	Conservative	0	Mismatches	72	Indels	1	Gaps	1	
Qy	88	CCCCTTCTACTGAAGCATG	CAGCGCCCTG	ACTCTCCGCA	CAC	TGGCCG	CCGAGGAC	GC	147	
Db	1	CCCCTCTCACAGAAGCAT	CAAGCGGTG	ACCTTACG	CAC	TGGCTGCC	GAGCAT	TT	60	
Qy	148	ACCAAGCATTACGAAC	TAGTCCGGGAG	CTGGGCAAGG	CACTATG	GGAAAG	GTTGAC	CTG	207	
Db	61	ACCAAGCATTATGACT	TGTGCGGAG	CTGGGCAAGG	CACTAC	GGGAAG	GTGAC	CTG	120	
Qy	208	GTGGTCTCAAGGGC	CACAGGCACAAAA	TGGC	ACTG	TAAGTTTGT	GAA	CAAGAGCAAA	ACC	267
Db	121	GTGGCTTACAGGGC	CACAGGCAC	TAAAA	TGGCCCTG	AAATTTGT	GAA	TAAAGTAAG	AACA	180
Qy	268	AAGCTGAAGAACTT	CTTACGGGAG	GTGAGCAT	CAACCA	ACAGCTCT	CTCTCC	AGCCCC	TTTC	327
Db	181	AAGCTGAAGAACTT	CTTCTGCGTGA	GGTGAGCAT	CAACCA	ACAGCTCT	GTGCT	TAGCCCC	TTTC	240
Qy	328	ATCATCAAGGTC	TTTGTGAGTGGT	CTTTGAG	CAGAGG	ACTGCTAC	GTCTTT	TGGCC	CAGAG	387
Db	241	ATCATCAAGGTC	TTTGTGAGTGGT	CTTTGAG	CCAGAG	GTGCTATG	CTTT	TGGCC	CAGAG	300
Qy	388	TACGCACCTGCTGGG	GCCTGTTTGA	CATCA	TCCTCC	CCAGGTGG	GGCTCC	CTGAG	GCAC	447
Db	301	TATGCACCTGCTGGG	GCCTGTTTGA	CATCA	TCCTCC	CCAGGTGG	GGCTCC	CAAG	GCAC	360
Qy	448	ACGGTGAAGCGCT	GTGTGCAG	CAGCTG	GGCCCTGG	CGCTGG	ACTCAT	GCACGGG	CGGCAG	507
Db	361	ACGGTGAAGCGCT	GTGTGCAG	CAGCTG	GGCGCTGG	CGCTGG	ACTCAT	GCACAGC	AGGCAG	420

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

Seq primer: pYX-5
 Location/Qualifiers
 1..719
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5703632"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP EWO"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GTGCGGGGAA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT

117 a 202 c 242 g 158 t
 Query Match 47.6%; Score 607; DB 13; Length 719;
 Best Local Similarity 90.3%; Pred. No. 1.2e-107;
 Matches 649; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

247 TTTCTGAACAGAGCAAAACCAAGCTGAAGCACTTCTTACGGAGGTGAGCATCACCAAC 306
 Db 1 TTTGTGAATAGATAGACAAAGCTGAGAACCTTCTGCGTGAAGTGAGCATCACCAAC 60
 307 AGCCTCTCTCCAGCCCTTCATCATCAAGTCTTTGACGFGTCTTTGACAGAGGAC 366
 Db 61 AGCCTGTGCTAGCCCTTCATCATCAAGTCTTTGACGFGTCTTCGAGACCGAGGAG 120
 367 TGCTACGCTTTTGCCAGGAGTACGACCTGCTGGGACCTGTTGACATCATCCTCC 426
 Db 121 TGCTATGCTTTGCCCAGGAGTATGACCTGCTGGGACCTGTTGACATCATCCTCC 180
 427 CAGTGGGGTCTCTGAGGACACGGTGAAGCGCTGTGTGACGACGTGGGCTGGCGCTG 486
 Db 181 CAGTGGGGTCTCCAGAGGACACGGTGAAGCGCTGTGTGACGACGTGGGCTGGCGCTG 240
 487 GACTTCATGACCGGCGGAGCTGTGACCGCGACATCAAGCCCGAGAGACGTGCTGCTG 546
 Db 241 GACTTCATGACACAGCGAGCTGTGACCGCGACATCAAGCCCGAGAGATGTGCTGCTG 300
 547 TTCACCGCGAGTCCCGCGGTAAAGCTGSCCGACCTTCGACGCGCGCTGAGGTGTGCCAGGCG 606
 Db 301 TTTGACCGTGAAGTCCCGCGGTGAAGCTGCTGCTGACATCAAGCCCGAGAGTGTGCCAGGCT 360
 607 TGCCGCGTCAAGCGCGTGAAGCGGACCACTCCCTTACACGCGCGCTGAGGTGTGCCAGGCG 666
 Db 361 TGCCGCTGAAGCGGTGAGCGGACATATACCGTACACAGCACCGGAGTGTGCCAGGCT 420
 667 GGCCGCCCGGACGGGCTGGCGGTGGACACGGGCGTGAAGTGTGGGCTTTCGGCGTCTC 726
 Db 421 GGCCGCCCGGATGCTTTCGGGCTGGACACGGGCGTGGACGCTGTGGGCTTTCGGCGTCTC 480

QY 727 ATCTTCTGCTGTACCGGCAACTTCCGTGGAGGCGGCGTCCGGCGCCGACGCTTC 786
 Db 481 ATCTTCTGCTGTACTGGCAACTTCCGTGGAGGCGTGCATCGGCTGCGGATGCTTC 540
 QY 787 TTCGAGGAGTTCGTGCGCTGCGACGCGGCGCTGCGCGGGCTGCTTCGAGTGGCGC 846
 Db 541 TTCGAGGAGTTCGTGCGCTGCGACGCGGCGTCCCTGCGGGGCTGCCATCGAGTGGCGC 600
 QY 847 CGCTTACCGAGCCCGCGCTGCGCATGTTCAGCGCTTACTTGGCCCTTGAGCCCGAGCGC 906
 Db 601 CGTTTACCGAGCGCTGCGTATGTTCAGCGGCTTCTGGCGCTAGAGCCTGAGCGG 660
 QY 907 CGCGGCCAGCAGGAGGTTCGCTTCTCTAAGCAGCAGCTCAGTCCGAGTGG 965
 Db 661 CGTGGGCCCGCCAGGAGGTCTTTCGCTTCTCTAAGCAGCAGCTCAGTCCGAGTGG 719

RESULT 11

BM944369
 LOCUS
 DEFINITION
 UI-M-EH0p-bvq-j-11-0-UI.r1 NIH BMAP_EH0p Mus musculus cDNA clone
 IMAGE:5695594 5', mRNA sequence.
 BM944369
 ACCESSION
 VERSION
 BM944369.1 GI:19427954
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 746)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-x@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES

Location/Qualifiers
 1..746
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5695594"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP_EH0p"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CAGCCAGAC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemlin Chin, Ph.D., program coordinator."
 127 a 206 c 251 g 161 t 1 others

BASE COUNT

127 a 206 c 251 g 161 t 1 others


```

351 QY CTTTGACACAGGAGCTGCTACGTCTTTGCCCCAGGAGTACGACCTGCTGGGGACCTGTT 410
    |||
188 Db CTTTGACACAGGAGGAGTGTATGCTTTGCCACAGGAGTATGACCTGCTGGGGACCTGTT 247
    |||
411 QY TGACATCATCCTCTCCCAAGGTGGGCTCCCTCAGAGACACGGTGAAGCGTGTGTGCAGCA 470
    |||
248 Db TGACATCATCCTCTCCCAAGGTGGGCTCCCAAGAGACACGGTGAAGCGTGTGTGCAGCA 307
    |||
471 QY GCTGGGCTGGCGCTGGAGCTTATGACACGGGGGAGCTGGTGACCCGCGCATCAAGCC 530
    |||
308 Db GCTGGGCTGGCGCTGGAGCTTATGACACAGCGAGCTGGTGACCCGCGCATCAAGCC 367
    |||
531 QY CGAGAACTGCTGCTGTTCCGACCGGAGTGGCGCGCTAAAGCTGGCCGACTTCGGCAT 590
    |||
368 Db CGAGAACTGCTGCTGTTGACCGTGAGTGGCGCGCTGTAAGCTGGCTGATTCGGCAT 427
    |||
591 QY GACGGCGCGCTGGCTGGCGGTGAAGCGGTGAGCGGCACATCCCTTACACGGCGCC 650
    |||
428 Db GACGGCGCGCTGGCTGGCGGTGAAGCGGTGAGCGGCACATCCCTTACACAGCACC 487
    |||
651 QY TCAGGTGTGCCAGGCGGGCGCGACACGGGCTGGCGGTGGACACGGGGGTGACGTGTG 710
    |||
488 Db CGAGGTGTGCCAGGCTGGCGCGGATGGCTTCGGCGGTGGACACGGGGGTGACGTGTG 547
    |||
711 QY GGCCTTGGCGCTGCTCATCTTCTGGGTGCTCACCGGCAACTTCCCGTGGGAGCGCGCTC 770
    |||
548 Db GGCCTTGGCGCTGCTCATCTTCTGGGTGCTCACCGGCAACTTCCCGTGGGAGCTGCTATC 607
    |||
771 QY GGGCGCGAGCGCTTCTTCGAGAGTTCGTGGCTGGAGCGGGCGCGCTCCCGGGGT 830
    |||
608 Db GGGTCGGATGCTTCTTCGAGAGTTCGTGGCTGGAGCGGGGTTCGGCTCCCGGGGT 667
    |||
831 QY GCTTCGCGAGTGGCGCGCTTACCGAGCGCGCGCTGGCGGTGGCGGTTCACGCGCTTACTGCG 890
    |||
668 Db GCGATCGAGTGGCGCGCTTACCGAGCGCGCGCTGGCGGTGGCGGTTCACGCGCTTACTGCG 727
    |||
891 QY CTTGAGCCCGAGCG 905
    |||
728 Db GCTAGAGCTGAGCG 742
    |||

RESULT 8
BM944510 758 bp mRNA linear EST 14-VAR-2002
LOCUS UI-M-BH0P-bvz-d-08-0-UI.r1 NIH_BMAP_EH0P Mus musculus cDNA clone
DEFINITION IMAGE:5695831 5', mRNA sequence.
ACCESSION BM944510
VERSION BM944510.1 GI:19428095
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 758)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..758
/organism="Mus musculus"

FEATURES
source

```

```

/mol_type="mRNA"
/strain="CS7BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5695831"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EH0P"
/organism="Mus musculus"
/vector="pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCAGGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT 108 a 229 c 265 g 153 t 3 others
ORIGIN
Query Match 48.0%; Score 612.6; DB 12; Length 758;
Best Local Similarity 88.5%; Pred. No. 9.7e-109;
Matches 676; Conservative 0; Mismatches 82; Indels 6; Gaps 1;
QY 328 ATCATCAAGGCTTTGACGTGCTTTGACGACAGGAGTCTGCTTTGCCAGGAG 387
Db 1 ATCATCAAGGCTTTGACGTGCTTTGACGACAGGAGTCTGCTTTGCCAGGAG 60
QY 388 TACGACCTGCTGGGGACCTGTTGACATCATCTCCCGAGGTGGGCTCCCTGAGAC 447
Db 61 TATGACCTGCTGGGGACCTGTTGACATCATCTCCCGAGGTGGGCTCCCTGAGAC 120
QY 448 ACGGTGAAGCGCTGTGTCAGCAGTGGGCTGCGGTGGACTTCAATGACAGGGCGGAG 507
Db 121 ACGGTGAAGCGCTGTGTCAGCAGTGGGCTGCGGTGGACTTCAATGACAGGGCGGAG 180
QY 508 CTGGTGACCGCGCATCAAGCCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 567
Db 181 CTGGTGACCGCGCATCAAGCCGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 568 GTAAAGCTGGCGGACTTCGGCATGACCGCGCGCTGGGCTGCCGTGCTGCTGCTGCTGCTG 627
Db 241 GTAAAGCTGGCGGACTTCGGCATGACCGCGCGCTGGGCTGCCGTGCTGCTGCTGCTGCTG 300
QY 628 GGCACCATCCCTTACAGCGCGCTGAGTGTGTCAGGCGGGCGCGCGCGCGCGCGCGCGCG 687
Db 301 GGCACCATCCCTTACAGCGCGCTGAGTGTGTCAGGCGGGCGCGCGCGCGCGCGCGCGCG 360
QY 688 GTGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 747
Db 361 GTGACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 748 AACTTCCCGTGGGAGGGCGGCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 807
Db 421 AACTTCCCGTGGGAGGGCGGCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 808 CAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 867
Db 481 CAGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 868 GGCATGTTCCAGCGCTTACTTGGCCCTGAGCCGAGCGCGCGCGCGCGCGCGCGCGCGCG 927
Db 541 GGCATGTTCCAGCGCTTACTTGGCCCTGAGCCGAGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 928 TTCGCGCTTCCCTCAAGCAGCAGCTCACGTCCGAGCTGCGCGCGCGCGCGCGCGCGCGCG 987
Db 601 TTCGCGCTTCCCTCAAGCAGCAGCTCACGTCCGAGCTGCGCGCGCGCGCGCGCGCGCGCG 660

```

oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GTCGCTGGGA. This library was created for the Project University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 108 a 243 c 275 g 158 t 2 others
 ORIGIN
 Query Match 48.6%; Score 619.2; DB 13; Length 786;
 Best Local Similarity 87.8%; Pred. No. 5.2e-110;
 Matches 688; Conservative 0; Mismatches 90; Indels 6; Gaps 1;
 QY 336 GGTCTTTGACGTGGTCTTTGACAGAGGACTGTACGCTTTTGCACGAGGTAGCGACC 395
 DB 1 GGTCTTTGACGTGGTCTTTGACAGAGGAGTGTATGCTTTTGCACGAGGTAGCGACC 60
 QY 396 TGCTGGGACCTGTTTGACATCATCCCTCCAGGTGGGGCTCCCTGAGGACACGGTGAA 455
 DB 61 TGCTGGGACCTGTTTGACATCATCCCTCCAGGTGGGGCTCCAGAGGACACGGTGAA 120
 QY 456 GCGCTGTGTGACAGCTGGGGCTGGCGCTGAGCTTATGACAGGGCGGCGAGTGTGCA 515
 DB 121 GCGCTGTGTGACAGCTGGGGCTGGCGCTGAGCTTATGACAGGAGCGAGTGTGCA 180
 QY 516 CCGGACATCAAGCCGAGAGAGCTGCTGTGTGACCGGAGTGGCGCGCTGAAGCT 575
 DB 181 CCGTACATCAAGCCGAGAGATGTGCTGTGTGACCGTGGTGGCGCGCTGAAGCT 240
 QY 576 GGCGACTTCGGATGACGCGCGCGTGGGTGCGCGCTCAAGCGCGTGAAGCGCACCAT 635
 DB 241 GGCTGACTTCGGATGACGCGCGCGTGGGTGCGCGCTGAAGCGCGTGAAGCGCACCAT 300
 QY 636 CCCTTACAGCGGCTGAGTGTGTCAGCGCGGCGCGCGCGCGCGCGTGGCGTGGACAC 695
 DB 301 ACCCTTACAGCAGCGGAGTGTGTCAGCGCTGGCGCGCGCGCGCTGGCGTGGACAC 360
 QY 696 GGGCGTGAAGCTGTGGGCTTTCGGCGTGTCTATCTTGTGCTGTCTACCGCACTTCCC 755
 DB 361 GGGGTGAGCTGTGGGCTTTCGGCGTGTCTATCTTGTGCTGTCTACCGCACTTCCC 420
 QY 756 GTGGAGCGCGCTGCGCGCGCGAGCTTCTTCAGAGAGTTCGTGCGCTGCGAGCGGG 815
 DB 421 GTGGAGCGCTGATCGGCTGCGAGTCCCTTCTTCAGAGAGTTCGTGCGCTGCGAGCGGG 480
 QY 816 CCGCTCCGGGCTGCTTTCGAGTGGCGGCTTACCGAGCGCGCTGCGCATGTT 875
 DB 481 TCGGCTCCGGGCTGCCATCGAGTGGCGCGCTTACCGAGCGCTGCGCATGTT 540
 QY 876 CCAGCGCTTACTGCGCTTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCGCTT 935
 DB 541 CCAGCGCTTCTGCGCTTAGAGCTGAGCGCGTGGCGCGCGCGCGCGCGCGCTTCCGCTT 600
 QY 936 CCTCAAGCAGAGCTCAGTCCGAGCTGCGCGCGCGCGCGCTTCCAGCGCGCGCGCGCGCG 995
 DB 601 CCTCAAGCAGAGCTCAGTCCGAGCTGCGCGCGCGCGCTTCCAGCGCGCGCGCGCGCG 660
 QY 996 CCCCGGAGCTCAA 1055
 DB 661 GCTTGGGACCGCTGC-----CAGGGTCCCTGCGCGCTTACCGCGCGCGCGCGCGCTCAA 714
 QY 1056 GCGGACGCTGTGACCGGAGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCTC 1115
 DB 715 GCGGACGCTGTGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 774
 QY 1116 GGTG 1119
 DB 775 CGTG 778

RESULT 7
 CA324262
 LOCUS
 DEFINITION
 ACCSSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CA324262 742 bp mRNA linear EST 26-NOV-2002
 UI-M-FY0-cco-1-22-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 IMAGE: 6822407 5', mRNA sequence.
 CA324262
 CA324262.1 GI:24542360
 EST.
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgbps-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

FEATURES
 source

Seq primer: pYX-5.
 Location/Qualifiers
 1..742
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone_lib="NIH_BMAP_FY0"
 /clone="IMAGE: 6822407"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH_BMAP_FY0"
 /notes="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCGAGACAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

BASE COUNT
 ORIGIN

128 a 199 c 255 g 159 t 1 others
 Query Match 48.4%; Score 617.2; DB 14; Length 742;
 Best Local Similarity 89.9%; Pred. No. 1.3e-109;
 Matches 661; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 171 GGAGCTGGGCAAGGACCTTATGGAGAGTTGACCTGGTGTGTACAGGGCACAGGCAC 230
 DB 8 GGAGCTGGGCAAGGACCTTATGGAGAGTTCGACCTGGTGTGTACAGGGCACAGGCAC 67
 QY 231 AAAATGGCACTGAAGTTTGTGAACAGACCAAAACCAAGCTCAAGAACTTCTCAGGGA 290
 DB 68 TAAATGGCCCTGAAATTTGTGAATAAGAGTAAGACAAAGCTGAAGAACTTTCTGCTGA 127
 QY 291 GGTGAGCATCAACCAAGAGCTCTCCAGCCCTTCATCATCAAGCTTTTACAGCTGGT 350
 DB 128 GGTGAGCATCAACCAAGAGCTCTCCAGCCCTTCATCATCAAGCTTTTACAGCTGGT 187

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 762)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..762
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="UI-M-FCO-cas-f-08-0-UI"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FCO"
 /note="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TGAGAGAGCC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 131 a 203 c 264 g 163 t 1 others
 ORIGIN

Query Match 48.6%; Score 619.2; DB 13; Length 762;
 Best Local Similarity 89.4%; Pred. No. 5.1e-110;
 Matches 666; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

170 GGGAGCTGGGCAAGGACCTATCGGAGGTTGACCTGGTGTACNAGGGACAGGCA 229
 Db 1 GGGAGCTGGGCAAGGACCTATCGGAGGTTGACCTGGTGTACNAGGGACAGGCA 60

230 CAAAATGGCCCTGAATTTGTGAATAGAGTAAAGCAAGCTGAAGAACTTCTACGGG 289
 Db 61 CAAAATGGCCCTGAATTTGTGAATAGAGTAAAGCAAGCTGAAGAACTTCTACGGG 120

290 AGGTGACATCACCAAGCTCTCTCCAGCCCTTCATCATCAAGTCTTTCAGCTGG 349
 Db 121 AGGTGACATCACCAAGCTCTCTCCAGCCCTTCATCATCAAGTCTTTCAGCTGG 180

350 TCTTTGAGACAGAGCTGTCTCTTTGCCCAGGAGTACGACCTGTGGGACCTGT 409
 Db 181 TCTTTGAGACAGAGCTGTCTCTTTGCCCAGGAGTACGACCTGTGGGACCTGT 240

410 TTGACATCATCCCTCCCGAGTGGGGTCCCTGTAGGACACGGTGAAGCGTGTGTGAGC 469
 Db 241 TTGACATCATCCCTCCCGAGTGGGGTCCCGAGAGGACACGGTGAAGCGTGTGTGAGC 300

470 AGCTGGGCTGGCGCTGGAATTATGACCGGGCGGAGCTGGTGCAACCGGACATCAAGC 529
 Db 301 AGCTGGGCTGGCGCTGGAATTATGACCGGGCGGAGCTGGTGCAACCGGACATCAAGC 360

Qy 530 CCGAGAACGTCTCTCTGTTGACCGGAGTCCCGCGGTAAAGCTGGCCGACTTGGCA 589
 Db 361 CCGAGAACGTCTCTCTGTTGACCGTGGTCCCGCGGTAAAGCTGGCTTGGCA 420

Qy 590 TGACGCGCGCGGTGGCTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGGC 649
 Db 421 TGACGCGCGCGGTGGCTGCGCGTCAAGCGCGTGAAGCGGACCATCCCTTACAGCGGC 480

Qy 650 CTGAGGTGTGTCAGCG 709
 Db 481 CCGAGGTGTGTCAGCG 540

Qy 710 GGGCTTTCGGCGTCTCATCTTCTGCGTCTACCGGCAACTTCCGTGGAGGCGCGGT 769
 Db 541 GGGCTTTCGGCGTCTCATCTTCTGCTGCTCACTGGCACTTCCGTGGAGGCTGCAT 600

Qy 770 CGGCG 829
 Db 601 CGGCG 660

Qy 830 TGCCTTCGAGTGGCG 889
 Db 661 TGCCTTCGAGTGGCG 720

Qy 890 CCCTGGAGCG 94
 Db 721 CGCTAGAGCTGAGCG 745

RESULT 6
 LOCUS BQ179876 786 bp mRNA linear EST 30-APR-2002
 DEFINITION UI-M-EWO-bww-a-08-0-UI.r1 NIH_BMAP_EWO Mus musculus cDNA clone
 IMAGE:5703823 5', mRNA sequence.
 ACCESSION BQ179876
 VERSION BQ179876.1 GI:20355368
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 786)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..786
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5703823"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EWO"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an

QY 481 CGCTGGACTTCATGACGCGCGGAGCTGTGTGACCGCGACATCAAGCCGAGAACTGT 540
 Db 497 CGCTGGACTTCATGACGAGGAGCTGTGTGACCGCGACATCAAGCCGAGAACTGT 556
 QY 541 CTGCTGTTTCGACCGGAGTGGCGCGTAAAGCTGGCGGACTTCGGCATCAGCGCGG 600
 Db 557 CTGCTGTTTCGACCGGAGTGGCGCGTAAAGCTGGCGGACTTCGGCATCAGCGCGG 616
 QY 601 GTGGGCTGCCGCTCAAGCGGCTGAGCGGACCATCCCTTACAGCGGCGCTGAGTGTGC 660
 Db 617 GTGGGCTGCCGCTGAGCGGCTGAGCGGACCATCCCTTACAGCGGCGCTGAGTGTGC 676
 QY 661 CAGCGGCGCGCGCGGAGTGGCGCGTAAAGCTGGCGGACTTCGGCATCAGCGCGG 720
 Db 677 CAGCGTGGCGCGCGGAGTGGCGCGTAAAGCTGGCGGACTTCGGCATCAGCGCGG 736
 QY 721 GTGCTCATCTTCTGCGTGTCTCAGCGG 746
 Db 737 GTGCTCATCTTCTGCGTGTCTCAGCTGG 762

RESULT 4
 BM964264
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BM964264 746 bp mRNA linear EST 18-MAR-2002
 UI-M-EQ0-bvu-m-02-0-UI.r1 NIH_BMAP_EQ0 Mus musculus cDNA clone
 IMAGE:5697193 5', mRNA sequence.
 BM964264
 BM964264.1 GI:19547684
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 746)
 NIH-MGC <http://hgnc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..746
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5697193"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EQ0"
 /note="Organ: brain; Vector: pYX-Aac; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Aac vector. The library tag
 sequence located between the Not I site and the polyA tail
 , is CAGCCACGAC. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous

FEATURES

source

System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."
 BASE COUNT 139 a 209 c 244 g 154 t
 ORIGIN
 Query Match 48.6%; Score 619.6; DB 12; Length 746;
 Best Local Similarity 89.4%; Pred. No. 4.3e-110;
 Matches 667; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 2 TGACGCTGGGCTGCCACAGAGCTGAGCGCGCCGCTCCCTGACCTGCTGTGGGCGGGGA 61
 Db 1 TGACGCTGGGCTGCCACAGAGCTGAGCGCGCCGCTCCCTGACCTGCTGTGGGCGGGG 60
 QY 62 CTGCCCCCTGGGCTGTGGCGGTGCCCCCTTCTCACTGAAGACATCAGCGCCCTGATCT 121
 Db 61 CCGCCCCCTTACAGAGTGCAGGTGTGCCCTCTCTACAGAAGACATCAAGCGCTGACCC 120
 QY 122 TCGCACACTGGCGCCGAGGAGCTCACCAAGCACTACGAACCTAGTCCGGSAGCTGGGCA 181
 Db 121 TACGCACACTGGCTGCCAGCGATGTTACCAAGCACTATGAGCTTGTCCGGAGCTGGGCA 180
 QY 182 AAGCACCTATGGAGAGGTTGACCTGTGTGTCTACAAGGGCACAGGCACAAAATGGCAC 241
 Db 181 AAGGACCTACGGGAGAGTGCACCTGTGTGTCTACAAGGGCACAGGCACAAAATGGCCC 240
 QY 242 TGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTACGGGAGGTGAGCATCA 301
 Db 241 TGAATTTTGTGAATAAGAGTAAGACAAAGCTGAAGAACTTCTGCGTGAAGTGAAGATCA 300
 QY 302 CCACAGCTCTCTCCAGCGCCCTTCACTATCATCAAGGTCTTTGAGTGGTCTTTGAGACAG 361
 Db 301 CCAACAGCTGTGTCTGAGCCCCCTTCACTATCAAGGTCTTTGAGTGGTCTTTGAGACAG 360
 QY 362 AGGACTGCTAGCTTTTGGCCAGAGTACGACCTGTCTGGGGACCTGTTTGACATCATCC 421
 Db 361 AGGAGTCTATGCTTTTGGCCAGAGTATGCACTGTCTGGGGACCTGTTTGACATCATCC 420
 QY 422 CTCGCCAGTGGGGCTCCCTGAGACACGCTGAAGCGCTGTGTGACAGAGTGGGCTGG 481
 Db 421 CTCCTCAGTGGGGCTCCCGAGAGACACGCTGAAGCGCTGTGTGACAGAGTGGGCTGG 480
 QY 482 CGCTGGACTTCATCACGGCGGCGAGCTGTGTGACCGCGACATCAAGCCGAGAACTGTC 541
 Db 481 CGCTGGACTTCATCACAGAGGCGAGCTGTGTGACCGCGACATCAAGCCGAGAACTGTC 540
 QY 542 TGTCTTTCGACCGGAGTGGCGCGCTAAAGCTGTGGCGACTTCGGCATCAGCGCGCG 601
 Db 541 TGTCTTTCGACCGTGAAGTGGCGCGCTGAAGCTGTGACTTCGGCATCAGCGCGCG 600
 QY 602 TGGCTCGCGGCTCAAGGGGTGAGCGGACCATCCCTTACAGCGGCGCTGAGGTGGCC 661
 Db 601 TGGGCTGCCGTGTGAAGCGGTGAGCGGCACTATACCGTACACAGCACCGGAGTGTGCC 660
 QY 662 AGCGGGCGCGCGCGAGCTGGCGGTGGACACGGCGGTGGAGCTGTGGGCTTTCGGCG 721
 Db 661 AGGTGGCGCGCGAGTGGCTTCGCGGTGGACACGGCGGTGGAGCTGTGGGCTTTCGGCG 720
 QY 722 TGCTCATCTTCTGCGTGTCTCAGCGG 747
 Db 721 TGCTCATCTTCTGCGTGTCTCAGCTGGC 746

RESULT 5
 BU611098
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BU611098 762 bp mRNA linear EST 20-FEB-2003
 UI-M-FC0-cas-f-08-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
 UI-M-FC0-cas-f-08-0-UI 5', mRNA sequence.
 BU611098
 BU611098.1 GI:23277313
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Qy 514 CACCGGACATCAAGCCCGAGAACGTGCTGCTGTTCACACCGAGTGGCGCCGCTAAAG 573
Db 64 CACCGGACATCAAGCCCGAGAACGTGCTGCTGTTCACACCGAGTGGCGCCGCTAAAG 123
Qy 574 CTGGCCGACTTCGGCATGACCGCGCGCTGGGCTGCGCGTCAAGCGGCTGAGCGGCAC 633
Db 124 CTGGCCGACTTCGGCATGACCGCGCGCTGGGCTGCGCGTCAAGCGGCTGAGCGGCAC 183
Qy 634 ATCCCTTACACGGCCCTGAGGTGTGCGACGGCGCGCGCCGCGAGCGGCTGGCGGTGAC 693
Db 184 ATCCCTTACACGGCGCTGAGGTGTGCGACGGCGCGCGCGCGCGAGCGGCTGGCGGTGAC 243
Qy 694 ACGGCGGTGAGCGTGTGGGCTTGGGCTGTCTATCTTCGTGCTGCTACCGGCACTTC 753
Db 244 ACGGCGGTGAGCGTGTGGGCTTGGGCTGTCTATCTTCGTGCTGCTACCGGCACTTC 303
Qy 754 CCGTGGGAGGCGCGCTCGGCGCGCGACCGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGG 813
Db 304 CCGTGGGAGGCGCGCTCGGCGCGCGACCGCTTCTTCGAGGAGTTCGTGCGTGGCAGCGG 363
Qy 814 GCGCGCTGCGGGGCTGCTTCGAGTGGCGCGCTTCACGAGCGCGCTGCGCATG 873
Db 364 GCGCGCTGCGGGGCTGCTTCGAGTGGCGCGCTTCACGAGCGCGCTGCGCATG 423
Qy 874 TTCAGCGCTTACTGGGCTGAGCGCGCGAGCGCGCGCGCGCTTCGAGGAGTTCGTGCG 933
Db 424 TTCAGCGCTTACTGGGCTGAGCGCGCGAGCGCGCGCGCGCTTCGAGGAGTTCGTGCG 483
Qy 934 TTCCTCAAGCAGAGCTCACGCTCGAGTGGCGCGCGCGCGCTTCGAGGAGTTCGTGCG 993
Db 484 TTCCTCAAGCAGAGCTCACGCTCGAGTGGCGCGCGCGCGCTTCGAGGAGTTCGTGCG 542
Qy 994 CCCCCGGGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1052
Db 543 CCCCCGGGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 601
Qy 1053 CAAGCGGAGGTGCTGACCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1111
Db 602 CAAGCGGAGGTGCTGACCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 661
Qy 1112 GGTGGTGGCTTGGCG 1168
Db 662 GGTGGTGGCTTGGCG 721
Qy 1169 CCGAG---CCCGGCTAGCTCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1225
Db 722 CCCAGAGACGAGCTAGACTCCCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 781
Qy 1226 AGAGCAAGGCGAGTGTGTGGC 1250
Db 782 AAAGAGCAAGGAGTGTGTGGC 806

RESULT 3
BU611087
LOCUS BU611087 762 bp mRNA linear EST 20-FEB-2003
DEFINITION UI-M-FCO-cas-b-20-0-UI.r1 NIH_BMAP_FCO Mus musculus cDNA clone
ACCESSION BU611087
VERSION BU611087
KEYWORDS EST
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 762)
TITLE NIH-MGC <http://mgi.nci.nih.gov/>
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5, Location/Qualifiers
1. 762
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-FCO-cas-b-20-0-UI"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FCO"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TGAGAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP); 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
BASE COUNT 145 a 214 c 248 g 154 t 1 others
ORIGIN

Query Match 48.6%; Score 620.2; DB 13; Length 762;
Best Local Similarity 89.4%; Pred. No. 3.3e-110;
Matches 667; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Qy 1 ATGAGCGTGGGCTGCCAGAGCGCTGAGCGCGCGCGCTCCCTGACCTGCTGTGGCGCGGG 60
Db 17 ATGAGCGTGGGCTGCCCTGAGCGCTGAAACCGCTCCACTCCCTGCTGCTGTGGCGCGGG 76
Qy 61 ACTGCCCTGGCGCTGTGCGCGGTGTGCGCTTCTCACTGAAGACATGCAAGCGCTGACT 120
Db 77 GCGCGCGCTGTACCGAGTGCAGGTGTGCGCTCCTCCTCAGAGACATGCAAGCGCTGACC 136
Qy 121 CTCGCGACACTGGCGCGCGAGCGTCAACAGCACTACGAACCTAGTCCGGAGCTGGC 180
Db 137 CTACGCGACCTGGCTGCCAGCGATGTTACCAAGCACTATGAGCTGTGTGGCGAGCTGGC 196
Qy 181 AAAGCGACCTATGGGAAGTTGACCTGTGTGTCTACAAGGCGCAGGCACAAAAATGGCA 240
Db 197 AAAGCGACCTATGGGAAGTTGACCTGTGTGTCTACAAGGCGCAGGCACCTAAATGGCC 256
Qy 241 CTGAAGTTTGTGAACAAGAGCAAAACCAAGCTGAAGAACTTCTTACCGGAGGTGAGCATC 300
Db 257 CTGAATAATTGTGAATAAGAGTAAGACAAAGCTGAAGAACTTCTTCTCGGTGAGGTGAGCATC 316
Qy 301 ACCAAGCGCTCTCTCCAGCGCTTTCATCATCATCAAGGTCTTTCAGCTGTGTGAGACA 360
Db 317 ACCAAGCGCTCTCTCTAGCCCTTTCATCATCAAGGTCTTTCAGCTGTGTGAGACA 376
Qy 361 GAGGACTGTCTACGCTTTCCTGGCGAGGAGTACGACCTGCTGGGAGCTGTGTGAGCATCATC 420
Db 377 GAGGAGTGTCTATGTCTTTCCTGGCGAGGAGTATGACCTGCTGGGAGCTGTGTGAGCATCATC 436
Qy 421 CCTCCCGAGGTGGGCGCTCCCTTGAGGACACCGGTGAAGCGCTGTGTGAGCAGCTGGCGCTG 480
Db 437 CCTCCTCAGGTGGGCGCTCCCGAGGAGCACCGGTGAAGCGCTGTGTGAGCAGCTGGCGCTG 496

Seq primer: pYX-5.
 Location/Qualifiers
 1. 804
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6402026"
 /tissue_type="whole brain"
 /dev_host="embryo 12.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP FCO"
 /note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; the library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TCAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
 BASE COUNT 142 a 221 c 272 g 169 t
 ORIGIN

Query Match 52.6%; Score 671.2; DB 13; Length 804;
 Best Local Similarity 89.7%; Pred. No. 4.4e-120;
 Matches 721; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

117 GACTCTCCGACACTGGCCGCGCAGCGAGTCACCAAGCACTAGCAACTAGTCGGGAGCT 176
 Db 1 GACCTACGCACACTGGCTGCCAGGATGTTACCAAGCACTATGAGCTTGTGGGAGCT 60

177 GGGCAAGGCACCTATGGGAAGTTGACCTGGTGTCTCAAGGGCACAGGCAAAAT 236
 Db 61 GGGCAAGGCACCTACGGGAAGTTCGACCTGGTGGCTTACAAGGGCACAGGCAAAAT 120

237 GGCACTGAAGTTGTGAACAGAGCAAAACCAAGCTGAAGCACTTCTACGGAGGTGAG 296
 Db 121 GGCCCTGAATTTGTGATTAAGAGTAAGAACAAGCTGAAGCACTTCTGGTGAGGTGAG 180

297 CATCACCAACAGGCTCTCCCTCCAGCCCTTCATCATCAAGGTCTTTGACGTGCTTTGA 356
 Db 181 CATCACCAACAGGCTCTCGCTAGCCCTTCATCATCAAGGTCTTTGACGTGCTTCGA 240

357 GACAGAGGACTGCTACTCTTTGCCAGGAGTAGCACTGCTGGGACCTGTTTGACAT 416
 Db 241 GACCGAGGAGTGCTATGCTTTGCCAGGAGTAGCACTGCTGGGACCTGTTTGACAT 300

417 CATCCCTCCCAAGTGGGGCTCCCTGAGGACACAGGTGAAGCGTGTGTGACAGAGCTGGG 476
 Db 301 CATCCCTCCCAAGTGGGGCTCCCAAGGACACAGGTGAAGCGTGTGTGACAGAGCTGGG 360

477 CTGGGGCTGGAATTCATGCAAGGGGGAGCTGGTGTGACCGGACATCAAGCCCGAGAA 536
 Db 361 GCTGGGGCTGGAATTCATGCAAGGAGGAGCTGGTGTGACCGTGAATCAAGCCCGAGAA 420

537 CGTGTGCTCTTCGACCGGAGTGCCCGCGGTAAAGCTGGCGACTTCGGCATGACGG 596
 Db 421 TGTGTCTCTCTTTGACCGTGAAGTGGCGCGGTGAAGCTGGCTGGCTGGCATGACGG 480

597 CCGCGTGGGCTCGCGCTCAAGGCGCTGAGCGGACCATCCCTACACGGCGCTGAGGT 656
 Db 481 CGCGTGGGCTCGCGCTGAGGCGCTGAGCGGACCATCCCTACACGACACCGGAGGT 540

657 GTGCCAGGCGGCGCGCGGCTGGCGGTGGCGGTGGACGTGGAGCGGTGGCGCTT 716
 Db 541 GTGCCAGGCTGGCGGCGCGGATGGCTTCGCGGTGGACACAGGCGGTGGAGCTTGGGCTT 600

QY 717 CGCGGTGCTCATCTTCTGGTGTCTACCGGCAACTTCCCGTGGGAGGCGGCGCTGCGGCGC 776
 Db 601 TGCGGTGCTCATCTTCTGGTGTCTACCTGCAACTTCCCGTGGGAGGCTGCACTCGGCTGC 660

QY 777 CGAGCGCTTCTTCGAGGAGTTGTGCGTGGCGAGCGGCGCGCTGCGGCGCTGCGCTTC 836
 Db 661 GGATGCTTCTTCGAGGAGTTGTGCGTGGCGAGCGGCGGTGCGCTGCGGCGCTGCGCATC 720

QY 837 GCAGTGGCGCGCTTCACCGAGCCCGCTGGCGATGTTCCAGCGCTTACTGCGCTGGA 896
 Db 721 CGAGTGGCGCGCTTACCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 897 GCCGAGCGCGCGCGCGCGCGCA 920
 Db 781 GCCTGAGCGCGTGGCGCGCGCA 804

RESULT 2
 EG683070
 LOCUS 602651746F1 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:4761754 5',
 DEFINITION mRNA sequence.
 ACCESSION EG683070
 VERSION EG683070.1 GI:13914467
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 888)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI614 row: 1 column: 11
 High quality sequence stop: 682.

FEATURES
 Location/Qualifiers
 1. 888
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4761754"
 /tissue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 47"
 /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACAGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC library."
 BASE COUNT 128 a 312 c 325 g 123 t
 ORIGIN

Query Match 49.5%; Score 630.6; DB 10; Length 888;
 Best Local Similarity 91.4%; Pred. No. 3.3e-112;
 Matches 736; Conservative 0; Mismatches 59; Indels 10; Gaps 6;

QY 454 AAGCGTGTGTGACGAGCTGGCGCTGGGCTGGGACTTCATGACGGGCGGAGCTGGTG 513
 Db 4 ATGCGCTGTGTGACGAGCTGGCGCTGGGCTGGGACTTCATGACGGGCGGAGCTGGTG 63

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2003, 15:05:53 ; Search time 2025 Seconds

(without alignments)
15302.818 Million cell updates/sec

Title: US-10-016-985-1

Perfect score: 1275

Sequence: 1 atgagcgtggcgtccagca.....ccatcgagatctgcgtctga 1275

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_esthm:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_estl:*

10: gb_estc:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vri:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	671.2	52.6	804	13	BU053381 UI-M-FCO-
2	630.6	49.5	888	10	BG683070 602651746
3	620.2	48.6	762	13	BU611087 UI-M-FCO-
4	619.6	48.6	746	12	BM964264 UI-M-EQO-

5	619.2	48.6	762	13	BU611098	BU611098 UI-M-FCO-
6	619.2	48.6	786	13	BQ179876	BQ179876 UI-M-EWO-
7	617.2	48.4	742	14	CA324262	CA324262 UI-M-FYO-
8	612.6	48.0	758	12	BM944510	BM944510 UI-M-EHOP
9	609.2	47.8	833	14	CA450675	CA450675 UI-M-FWO-
10	607	47.6	719	13	BQ745423	BQ745423 UI-M-EWO-
11	606.8	47.6	746	12	BM944369	BM944369 UI-M-EHOP
12	603.8	47.4	730	13	BU701617	BU701617 UI-M-FYO-
13	603.2	47.3	733	12	BM947152	BM947152 UI-M-EHOP
14	596.8	46.8	706	13	BQ179888	BQ179888 UI-M-EWO-
15	594.4	46.6	767	13	BU613351	BU613351 UI-M-EWO-
16	591.8	46.4	706	13	BQ745569	BQ745569 UI-M-EWO-
17	591.4	46.4	764	13	BQ180145	BQ180145 UI-M-EWO-
18	589.8	46.3	731	14	CA317985	CA317985 UI-M-FWO-
19	587.6	46.1	730	13	BU613504	BU613504 UI-M-EWO-
20	587.4	46.1	693	14	CA316498	CA316498 UI-M-FWO-
21	587.2	46.1	696	12	BM951369	BM951369 UI-M-EHO-
22	585.8	45.9	733	12	BM944577	BM944577 UI-M-EHOP
23	583.4	45.8	748	14	CA317267	CA317267 UI-M-FWO-
24	582.6	45.7	698	12	BM950544	BM950544 UI-M-EHOP
25	582	45.6	686	13	BU708991	BU708991 UI-M-EWO-
26	579.2	45.4	720	14	CB526693	CB526693 UI-M-FYO-
27	578.6	45.4	700	12	BM944032	BM944032 UI-M-EHOP
28	577.8	45.3	732	14	CA319385	CA319385 UI-M-FWO-
29	576	45.2	713	14	CA318039	CA318039 UI-M-FWO-
30	571	44.8	684	13	BQ768987	BQ768987 UI-M-FCO-
31	569.6	44.7	673	12	BM950706	BM950706 UI-M-EHOP
32	568.6	44.6	736	14	CA319869	CA319869 UI-M-FWO-
33	567.6	44.5	692	14	CA318185	CA318185 UI-M-FYO-
34	557.4	43.7	687	14	CB527765	CB527765 UI-M-FYO-
35	547.4	42.9	675	12	BM944210	BM944210 UI-M-EHOP
36	545.8	42.8	716	14	CA325404	CA325404 UI-M-FYO-
37	541.6	42.5	655	14	CA316925	CA316925 UI-M-FWO-
38	539.4	42.3	736	14	CA328687	CA328687 UI-M-FYO-
39	537.6	42.2	777	12	BM947305	BM947305 UI-M-EHOP
40	536.6	42.1	698	14	CB249777	CB249777 UI-M-EXO-
41	534.6	41.9	629	13	BU709002	BU709002 UI-M-EWO-
42	531.2	41.7	677	12	BM946930	BM946930 UI-M-EHOP
43	526.4	41.3	723	14	CA315154	CA315154 UI-M-FWO-
44	525.2	41.2	684	13	BU612423	BU612423 UI-M-EWO-
45	523.2	41.0	716	13	BU053232	BU053232 UI-M-FCO-

ALIGNMENTS

RESULT 1
BU053381
LOCUS
DEFINITION
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EWAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.
BU053381
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU053381 804 bp mRNA linear EST 26-AUG-2002
UI-M-FCO-bxd-e-03-0-UI.r1 NIH_EWAP_FCO Mus musculus cDNA clone
IMAGE:6402026 5', mRNA sequence.
BU053381
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

```

Db          531 EG 532
RESULT 15
ULK1 MOUSE
ID ULK1_MOUSE STANDARD; PRT: 1051 AA.
AC 070405;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase ULK1 (EC 2.7.1.-) (Unc-51-like kinase
DE 1) (Unc51.1 serine/threonine kinase).
GN ULK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=26262945; PubMed=9600096;
RX Yan J., Kuroyanagi H., Kuroiwa A., Matsuda Y., Tokumitsu H.,
RA Tomoda T., Shirasawa T., Muramatsu M.-A.;
RT "Identification of mouse ULK1, a novel protein kinase structurally
RT related to C. elegans UNC-51.";
RL Biochem. Biophys. Res. Commun. 246:222-227(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=20088285; PubMed=10624947;
RA Tomoda T., Shatt R.S., Kuroyanagi H., Shirasawa T., Hatten M.E.;
RT "A mouse serine/threonine kinase homologous to C. elegans UNC51
RT functions in parallel fiber formation of cerebellar granule
RT neurons.";
RL Neuron 24:833-846(1999).
RC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC AFG1/UNC-51/ULK1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF053756; AAC40118.1; -.
CC FIRM; AF072370; AAF23317.1; -.
CC HSP; P24941; ICKP.
CC MGD; MGI:1270126; Ulk1.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF001245; Tyr_pkinase.
CC PRINTS; PR001069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 16 278 PROTEIN KINASE.
FT NP BIND 22 30 ATP (BY SIMILARITY).
FT BINDING 46 46 ATP (BY SIMILARITY).
FT ACT SITE 138 138 BY SIMILARITY.
FT DOMAIN 297 310 POLY-SER.
SQ SEQUENCE 1051 AA; 112462 MW; 998021985FB4E8A0 CRC64;

```

Query Match 13.9%; Score 311.5; DB 1; Length 1051;
 Best Local Similarity 25.6%; Pred. No. 4.3e-12;
 Matches 127; Conservative 60; Mismatches 146; Indels 163; Gaps 24;

```

QY 56 VRELGRGTGKVDL-----VYKGT-----GTRMALKFVNKSK-TLKNFL-REVSIT 101
Db 8 VETVGKFEFSKDLIGHGAFVVFVKGREREKHDLEVAVKCINKNLAQSTLLGKEIKIL 67
QY 102 NSLSSSPFIKVEDVVFETEDCVFAQAYAPAGDLFDIIPQVGLPDDITVKRCVQOGLA 161
Db 68 KELKHEN-IVALYDFQEMANSVYL-VMEYCNGGDLADYLHMTLSEDTVRLFLQOIAGA 125
QY 162 LDFMGRQLVHRDIKPEVILLFDRECR-----RVKLADFGMTRVGCVRKVS--GTI 212
Db 126 MRLHLSKGIHRLDKPQNILSNPGRBRANPSNIRVKIADFGFARYLQSNMMAATLCGSP 185
QY 213 PYTAPEVCOAGRADGLAVDTGVDVNAFGLVIFCVLTGNFPEWAAAGAD--AFPEEFVRWQ 270
Db 186 MYMAPEVINSQHYDGKA-----DLWSIGITIVYQCLTGKAPFQASSPQDLRLFYEK---N 236
QY 271 RGLPLGLPSQWRRTFEPALRMFORILALEPERRGPAKEVFRFLKH---ELTSELRRRP-- 325
Db 237 KTLVPAIP-----RETSAPLR--QLLLALLQRNHKDRMDFEFHFFHFLDASTPIKSPPV 290
QY 326 -----SHRARKPP-GDRP-----PAAGPLR----- 344
Db 291 PVESYPSGSGSSSSSSSSSHLASPSPJGEMPOLQKLTSPADAGFLQSRDSGSSSKD 350
QY 345 -----LEAPGLKRTV-----LTPESGS-----GSRPAP-PAVGS 372
Db 351 SCDTDDFVMVPAQFPGLVAEASAKPFPDSLCSGSSSLVASAGLESHGRTSPSPSCSS 410
QY 373 VPLPVVPVVP-----VPVPEVP-----VPEP----- 392
Db 411 SPSPSGRPGPFSSNRYGASVPIPVPTQVHNYQRIEQNLQSPQOQTARSSAIRRSGSTTP 470
QY 393 ---GLAPQGPFGRTDG 405
Db 471 LGFGRASPSPPSHTDG 486

```

Search completed: December 5, 2003, 09:56:51
 Job time : 18 secs

```
CC CC -!- FUNCTION: CONTROLS CELL CYCLE CHECKPOINT. MAY PARTICIPATE IN
CC CC TRANSITION OF THE DNA DAMAGE AND REPLICATIONAL STRESS SIGNALS.
CC CC INHIBITS CDC25 PHOSPHATASE BY PHOSPHORYLATING IT, PREVENTING THE
CC CC ENTRY INTO MITOSIS. MAY HAVE A ROLE IN MEIOSIS AS WELL.
CC CC -!- ENZYME REGULATION: KINASE ACTIVITY IS UPREGULATED BY
CC CC AUTOPHOSPHORYLATION. RAPIDLY PHOSPHORYLATED IN RESPONSE TO DNA
CC CC DAMAGE AND TO REPLICATION BLOCK (BY SIMILARITY).
CC CC -!- SUBCELLULAR LOCATION: Nuclear
CC CC -!- SIMILARITY: BELONGS TO THE SER/THER FAMILY OF PROTEIN KINASES.
CC CC CDS1 SUBFAMILY.
CC CC -!- SIMILARITY: Contains 1 FHA domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch)
CC CC -----
CC CC EMBL; AF086905; AAC83694.1; -.
CC CC HSSP; Q63450; IA06.
CC CC MGD; MG1:I355321; Chek2.
CC CC InterPro; IPR000253; FHA.
CC CC InterPro; IPR000719; Prot_kinase.
CC CC InterPro; IPR002290; Ser_thr_pkinase.
CC CC Pfam; PF00498; FHA; 1.
CC CC Pfam; PF00069; pkinase; 1.
CC CC ProDom; PD000001; Prot_kinase; 1.
CC CC SMART; SM00240; FHA; 1.
CC CC SMART; SM00220; S_TKC; 1.
CC CC PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC CC PROSITE; PS50006; FHA_DOMAIN; 1.
CC CC Transferase; Serine/threonine-protein kinase; ATP-binding;
CC CC Cell cycle; Phosphorylation; Nuclear protein.
CC CC 117 179
CC CC DOMAIN 179
CC CC 224 490 PROTEIN_KINASE.
CC CC NP_BIND 230 238 ATP (BY SIMILARITY).
CC CC BINDING 253 253 ATP (BY SIMILARITY).
CC CC ACT_SITE 351 351
CC CC SEQUENCE 546 AA; 61088 MW; A7949EFB5572CDA CRC64;
```

Query Match	14.0%; Score 314.5; DB 1; Length 546;
Best Local Similarity	27.1%; Pred. NO. 1.4e-12;
Matches	98; Conservative 67; Mismatches 152; Indels 45; Gaps 12;
QY	24 GPGAGVPLLTEDMQALTL-----RTLAASDVTKHYELVLRELKGTGYKVGD 68
DB	180 GKXKRCPLSNNSEIALSLCNKVFVEFLTVDDQSYYVKPELRDEYIMSKTLLSGACGEVK 239
QY	69 LVVYKGTGTFMALKFVNKSKTKL-----KNFLREVSITNSLSSSPFIILKVFVDVFE 119
DB	240 MAFERKTCOKVAIKISKRRFALGSSREADTAPSVETEIEILKL-NHPCIIKIKD-VFD 297
QY	120 TEDCVYFAOEYPAGDGLFDIIPQGVLPBDTVKRCVQQLGLADFMHGROLVHRDIKEN 179
DB	298 AEDYIT-VLELMEGGELEFRVGNKKLKCATCKLYFYOMLVAVQYLHENGIILHDLKPEN 356
QY	180 VLLFDRE--CRRVKLADFQMTRRVG--CRVKRVSIGTIPTAPEVCOCAGRADGLAVDTGVD 235
DB	357 VLLSSGEEDC-LIKITDFQGSKILGETSLMRTLCGITYLAPELVLSNGTAGYS--RAVD 413
QY	236 VNAGVLLIFCVLTGNFPWRAASGADAFFEEFVFWQGRRLPGLFSQWRRFTEPALRMFORL 295
DB	414 CWSLGVLFTICLSGYPFSEHKTVQSIKDQIT---SGKYNFIPFWMTVDVSEBALDIVVKKL 470
QY	296 LALEPERRGPAKEVER--FLKHETLSE-----LRRRPSHRANKPPGDRPPAAGPLREA 347
DB	471 LVDVPKARLTTEALNHPMLQDEYMKKKFODLLVQEKNSVTLPVAPAQTSSQKPLELEV 530
QY	348 PG 349

```

RN  SEQUENCE FROM N.A.
RP  TISSUE=Brain;
RX  MEDLINE=98290545; PubMed=9628591;
RA  Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA  Nomura N., Ohara O.;
RT  "Prediction of the coding sequences of unidentified human genes. IX.
RT  The complete sequences of 100 new cDNA clones from brain which can
RT  code for large proteins in vitro.";
RL  DNA Res. 5:31-39(1998).
CC  -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
DR  EMBL; AB011109; BAA25463.1; -.
DR  HSRF; F24941; IBUH.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_kinase.
DR  InterPro; IPR001245; Tyr_pkinase.
DR  Pfam; PF00069; pkinase_1.
DR  PRINTS; PR00109; TYRKINASE.
DR  ProDom; PD000001; Prot_kinase; 1.
DR  SMART; SM00220; S_TKc; 1.
DR  PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR  PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR  PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW  Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW  ATP-binding.
FT  DOMAIN 55 306 PROTEIN KINASE.
FT  NP_BIND 61 69 ATP (BY SIMILARITY).
FT  BINDING 84 84 ATP (BY SIMILARITY).
FT  ACT_SITE 178 178 BY SIMILARITY.
SQ  SEQUENCE 661 AA; 74304 MW; 806F37D52CA4710F CRC64;

Query Match 14.1%; Score 316.5; DB 1; Length 661;
Best Local Similarity 27.5%; Pred. No. 1.3e-12;
Matches 110; Conservative 65; Mismatches 156; Indels 69; Gaps 17;

Qy  20 GTAPGPGAGVLLLEDQQA-----LILRTLAASDVTKH-----YELVRLGKGTGKV 67
Db  10 GDRPDLGLGAPSPREAVAGATAALEPRKPGVKRHHKHNKHYELQETLGRGTGKV 69
Qy  68 DLVYKGTGTQWALKFVNKSKTK---LKNFLREVSTINSLSPPFIKVPDVFETEDC 123
Db  70 KRATERFSGRVAIKSRKDKIKDEQDMVHRRRIEMSSL-NHPHLSIYE-VFENKDK 127
Qy  124 YVFAQEVAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRLVHRDIKPENVLLF 183
Db  128 IVIIMEYASKGELYDISERRLSERTRHFRFOIVSAVHYCHKNVGVHRDLKENILL- 186
Qy  184 DRECRVKLADFGMTR--RVGCRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVDVAFGV 241
Db  187 DDNC-NIKIADFGLSNLYQKDKFLQTCGSPLYASPEIVNGRPYRGPEVDS---WALGV 241
Qy  242 LIFCVLTGTFWEAASGADAFEFBFVWQGRGLPGLPSQWRRTFEP--ALRMFORLLALE 299
Db  242 LLYTLVVTWTFD---GFD--HKNLIRQISS-----GEYREPTQPSDARGLLIRMLWN 290
Qy  300 PERGPAKEVRF-----LKHSLTSELRERP---SHRA-----RKPP 333
Db  291 PDRATTIEDIANHWVNWGYKSSVCCDCLDHSSEPLARIIDWHHRSTGLQADTEAKM 350
Qy  334 GDRPPAGPLRLLEAPGLKRTV-----LTSGSGSRPAPPA 369
Db  351 GLAKPTTSEVMLERQRLSKKSKENDFAQSQDQAVPESPS 390

```

```

RESULT 13
AAK2_RAT
ID AAK2_RAT STANDARD; PRT; 552 AA.
AC Q09137;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, catalytic alpha-2 chain (EC 2.7.1.-)
DE (AMPK alpha-2 chain)
GN PRKA2 OR AMPK2 OR AMPK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94209324; PubMed=7908907;
RA Carling D., Aguan K., Woods A., Verhoeven A.J.M., Beri R.K.,
RA Brennan C.H., Sidebottom C., Davison M.D., Scott J.;
RT "Mammalian AMP-activated protein kinase is homologous to yeast and
RT plant protein kinases involved in the regulation of carbon
RT metabolism.";
RL J. Biol. Chem. 269:11442-11448(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95234757; PubMed=7718624;
RA Gao G., Widmer J., Stapleton D., Teh T., Cox T., Kemp B.E.,
RA Witters L.A.;
RT "Catalytic subunits of the porcine and rat 5'-AMP-activated protein
RT kinase are members of the SNF1 protein kinase family.";
RL Biochim. Biophys. Acta 1266:73-82(1995).
CC -!- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
CC BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
CC CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
CC HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGLUTARYL-COA REDUCTASE.
CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
CC SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
CC DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
CC AND/OR HYPOXIA. THIS IS A CATALYTIC SUBUNIT.
CC -!- SUBUNIT: HETEROTRIMER OF A CATALYTIC SUBUNIT, A BETA AND A GAMMA
CC NON-CATALYTIC SUBUNITS.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q09137-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q09137-2; Sequence=VSP_004949, VSP_004950;
CC Note=Lacks the sequence parts essential for kinase activity and
CC is therefore inactive;
CC -!- TISSUE SPECIFICITY: SKELETAL MUSCLE, LOWER LEVELS IN LIVER, HEART,
CC AND KIDNEY.
CC -!- INDUCTION: BY AMP.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. SNF1
CC SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; Z29486; CAA82620.1; -.
DR  EMBL; U12149; AAB5033.1; -.
DR  PIR; A53621; A53621.
DR  HSSP; Q63450; 1A06.
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_kinase.
DR  InterPro; IPR001245; Tyr_pkinase.
DR  Pfam; PF00069; pkinase; 1.

```


RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN [7]
RP SEQUENCE OF 323-775 FROM N.A. (ISOFORM 7).
RC TISSUE=Urinary bladder;
RC Reynolds C.H., Patel U.A., Anderton B.H.;
RA "Homo sapiens mRNA partial sequence for a protein kinase, STK10,
RT similar to p78/C-TAK1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in the specific phosphorylation of microtubule-
CC associated proteins for tau, MAP2 and MAP4. Phosphorylated Cdc25C
CC on Ser-216
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=7;
CC Name=1;
CC IsoId=P27448-1; Sequence=Displayed;
CC Name=2; Synonyms=CTAK75a;
CC IsoId=P27448-2; Sequence=VSP_004944;
CC Name=3;
CC IsoId=P27448-3; Sequence=VSP_004942, VSP_004944;
CC Name=4;
CC IsoId=P27448-4; Sequence=VSP_004942, VSP_004945;
CC Name=5;
CC IsoId=P27448-5; Sequence=VSP_004942;
CC Name=6; Synonyms=p58;
CC IsoId=P27448-6; Sequence=VSP_004942, VSP_004943, VSP_004944;
CC Name=7;
CC IsoId=P27448-7; Sequence=VSP_004943;
CC -!- TISSUE SPECIFICITY: Ubiquitous
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MARK
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 1 kinase-associated (KAI) domain.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U64205; AAC15093.1; --
CC EMBL; AF159295; AAD48007.1; --
CC EMBL; AF387637; AAK82367.1; --
CC EMBL; AF465413; AAL69982.1; --
CC EMBL; M80359; AAA59991.1; --
CC EMBL; BC024773; AAH24773.1; --
CC EMBL; AF170723; AAD51631.1; --
CC HSSP; Q63450; IAO6.
CC Genew; HGNC:6897; MARK3.
CC MIN; 602678; --
CC InterPro; IPR001772; Kinase Cterm.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR000449; UBA_domain.
CC Pfam; PF02149; KAI; 1.
CC Pfam; PF00659; pkinase; 1.
CC Pfam; PF00627; UBA; 1.
CC PRINTS; PR00109; TYRKINASE.
CC PRODOM; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC SMART; SM00165; UBA; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50030; UBA; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 56 330 PROTEIN KINASE.
FT DOMAIN 326 365 UBA.
FT NP_BIND 62 70 ATP (BY SIMILARITY).
FT BINDING 85 85 ATP (BY SIMILARITY).
FT ACT_SITE 201 201 BY SIMILARITY.
FT VARSPPLIC 162 184 Missing (in isoform 3, isoform 4, isoform
FT 5 and isoform 6).
FT FTId=VSP_004942.
FT VARSPPLIC 394 409 Missing (in isoform 6 and isoform 7).
FT VARSPPLIC 638 661 Missing (in isoform 2, isoform 3 and
FT isoform 6).
FT FTId=VSP_004944.
FT VARSPPLIC 638 646 Missing (in isoform 4).
FT CONFLICT 125 125 E -> Q (IN REF. 5).
FT CONFLICT 139 139 E -> K (IN REF. 4 AND 5).
FT CONFLICT 149 149 R -> K (IN REF. 2).
FT CONFLICT 448 448 A -> G (IN REF. 5).
FT CONFLICT 466 466 S -> G (IN REF. 1 AND 2).
FT CONFLICT 479 479 S -> T (IN REF. 5).
FT CONFLICT 539 539 A -> D (IN REF. 5).
FT CONFLICT 626 626 N -> T (IN REF. 5).
FT CONFLICT 668 668 E -> K (IN REF. 5).
SQ SEQUENCE 776 AA; 86944 MW; A245496849070098 CRC64;
Query Match 14.2%; Score 317; DB 1; Length 776;
Best Local Similarity 28.7%; Pred. NC. 1.4e-12;
Matches 91; Conservative 64; Mismatches 104; Indels 58; Gaps 13;
QY 37 QALTLTAT-----LAASDVTKH--YELVRELKGTGKVDLVVYKGTGTMALK 82
DB 26 QEVTSKTSRGACRNSIASCADEQPHIGNYRLKTTIGKGNFAKVLARHLTGREVALK 85
QY 83 FVNKSK---TKLKNFLREVSITNSLSSPFIKVFDFWFEFETEDCYVFAOEYAPAGDLFDI 139
DB 86 IIDKTQLNPTSLQKLFREVRIMKIL-NEFNIVKLFVEI-ETEKTLYLIMEYASGGEVFDY 143
QY 140 IPPQVGLPDETIVKRCVQQ-----LGL-----ALDFHMGRLVHRDIK 176
DB 144 LVAHGRMKKEARSKPRQCGAGQTIKQVSDLLSLMFTFIVSAVQYCHQKRIVRDLK 203
QY 177 PENVLFDRECRVRKLADFGMTR--VGCRVKRVSGTIPVTAPEVCQAGRADGLAVDTGV 234
DB 204 AENLLI-DAD-MNIKIADFGFSNEFTVGGKIDTFGSPPPYAAPELFGKKYDG---PEV 257
QY 235 DVWAFGLVIFCVLTGNFPWEAASGADAFPEEFVWRQGRILPGLPSQWRRTFEPALRMFOR 294
DB 258 DVWSGLVILTVSGSLPFDGQNLKE--LRERVLRGKYRIPFYMS-----TDCENLLKR 309
QY 295 LLALEPERRGPAKEVFR 311
DB 310 FLVINPITKGTLEQIMK 326
RESULT 12
Y537_HUMAN
ID Y537_HUMAN STANDARD; PRT; 661 AA.
AC O60285;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable serine/threonine-protein kinase KIAA0537 (EC 2.7.1.1-).
GN KIAA0537.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkert G., Aert R., Robben J., Grynoprez B.,
 RA Wältjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Ravaglia J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrall B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; AL031543; CA220833.1; -
 DR PIR; T41587; T41587.
 DR HSP; Q63450; IAO6.
 DR GeneDB SPombe; SPCC74.03c; -
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR000449; UBA_domain.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00030; UBA; 1.
 DR Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 34 285 PROTEIN KINASE.
 FT DOMAIN 304 345 UBA.
 FT NP_BIND 40 48 ATP (BY SIMILARITY).
 FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT_SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 576 AA; 65996 MW; 55857E8F171E7B50 CRC64;
 Query Match 14.2%; Score 317.5; DB 1; Length 576;
 Best Local Similarity 30.8%; Pred. No. 9.5e-13;
 Matches 98; Conservative 56; Mismatches 101; Indels 63; Gaps 15;
 QY 42 RTLAASDVTHY----ELVRE-LGKGTGKVDLVVYKGTGTQKALKFNKSKTKLKNF--- 94
 DB 19 RVLPPEAIKSRHIGFYIIRETLGEGSGKVKVLAHYTKQKVALKFKISRLKKSDMHR 78
 QY 95 -LREVSITNSLSPPFIKVDVVNFEEDCVFAQEVAPAGDLIDIPQVGLPEDIIVKR 153
 DB 79 VEREISVL-KLLRHPIHIIKLYDVITPTTD-IWMVIEA-GELPDIYVEKKRMTEDEGR 135
 QY 154 CVOQLGLALDFPMHGRQLVHRDIKPEENVLLFDRECRVKLADFGMTRRV--GCRVKRVSGT 211
 DB 136 FFOQICAEIYCHRRKIVHRDLKENLLDDN--LNVKIADFGLSNMTDGNFKTSKGS 193
 QY 212 IPTAPEVCQGRADGLAVDTGVVDVAFVGLIFCVLTGNFPWEAASGADAFFEEFVRQR 271
 DB 194 PNYAAPEYING----KLYAGPEVDVMSGCVILYVLMVGRLP-----FDDEFI---- 236

QY 272 GRUGPGLPSQWR-----FTEP-ALRMFORLLALEPERRGPAKEVFRFLKHELT-S 319
 DB 237 -----PNLFKKVNSCVVVPDPLSPGASQLIRRMIVADPMQR-----ITIQ 277
 QY 320 ELRRRPSHRARKPPGDRP 337
 DB 278 EIRRDPMFNVNLDYLRP 295
 RESULT 11
 MRK3 HUMAN
 ID MRK3_HUMAN STANDARD; PRT; 776 AA.
 AC P27448; O60219; Q8TB41; Q8WK83; Q96RGI; Q9UMY5; Q9UN34;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE MAP/microtubule affinity-regulating kinase 3 (EC 2.7.1.27) (Cdc25C-
 DE associated protein kinase 1) (CTAK1) (C-TAK1) (Serine/threonine
 DE protein kinase p78) (Ser/Thr protein kinase PAR-1) (protein kinase
 DE STRK10).
 DE MARK3 OR CTAK1.
 GN Homo sapiens (Human).
 OS
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX MEDLINE=98202387; PubMed=9543386;
 RA Peng C.Y., Graves P.R., Ogg S., Thoma R.S., Byrnes M.J. III, Wu Z.,
 RA Stephenson M.T., Pionica-Worms H.;
 RA "C-TAK1 protein kinase phosphorylates human Cdc25C on serine 216 and
 RA promotes 14-3-3 protein binding.";
 RL Cell Growth Differ. 9:197-208(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Waggoner S.N., Stephen R., Farrar W.L., Howard O.M.Z.;
 RA "Human serine/threonine protein kinase cTAK1/Kp78/Mark3:
 RT Identification of a novel splice variant and a larger 5'UTR.";
 RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RX MEDLINE=21326471; PubMed=11433294;
 RA Sun T.Q., Lu B., Feng J.J., Reinhard C., Jan Y.N., Fantl W.J.,
 RA Williams L.T.;
 RA "PAR-1 is a Dishevelled-associated kinase and a positive regulator of
 RT Wnt signalling.";
 RL Nat. Cell Biol. 3:628-636(2001).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 5).
 RX Drewes G.;
 RA "Characterization of an alternatively spliced form of MARK3 from human
 RT brain.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 6).
 RA Maheshwari K.K., Som S., Parsa I.;
 RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Cervix;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.C., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

```

DR  Pfam: PF00069; Pkinase; 1.
DR  ProDom: PD000001; Prot_Kinase; 1.
DR  SMART: SM00220; S_TK; 1.
DR  PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR  PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR  PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR  Transferase: Serine/threonine-protein kinase; ATP-binding;
KW  Phosphorylation; Carbohydrate metabolism; Nuclear protein.
KW  DOMAIN 16 29 POLY-HIS.
FT  DOMAIN 53 305 PROTEIN_KINASE.
FT  NP_BIND 59 67 ATP (BY SIMILARITY).
FT  BINDING 82 82 ATP (BY SIMILARITY).
FT  ACT_SITE 175 175 BY SIMILARITY.
FT  MOD_RES 208 208 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT  CONFLICT 7 8 PQ -> AR (IN REF. 2).
FT  CONFLICT 228 234 AGEVDV -> SSVQKMI (IN REF. 2).
FT  CONFLICT 242 255 YVMLCGRLPFDDEF -> GMSCCVVDHSMST. (IN
FT  CONFLICT 242 255 RE 2).
FT  FT 387 387 A -> R (IN REF. 2).
FT  CONFLICT 416 416 R -> A (IN REF. 2).
FT  CONFLICT 494 494 S -> L (IN REF. 2).
FT  SEQUENCE 620 AA; 70005 MW; 1806C652B5061D2B CRC64;
SQ
Query Match 14.5%; Score 324; DB 1; Length 620;

```

QY	53	YELVRELGGTGYKQDLVYVKGTTGKALKFVNKSKTKLKNFL-----REVSTNSLSSS	107
Db	53	YQTLKTLGGSGFGKVLAKHLGTTGQGVALKIINR-KTAKSDMQGRVEREISYLR-LLRH	110
QY	108	PRIKVDFVVFETDCVFAQAEVAPAGDLFDIIPQVGLPDDTVKRCVQQLGLALDFMHG	167
Db	111	PHIKYDVI-KSKDIIWVIEFA-GKELFDYIVQRGKMEDEARFFQQLIAAVEYCHR	168
QY	168	ROLVHRDRIKPNVLLDFEDRCRRVKYKLDLFGMTRRV--GCRVKRVSGITPY-TAPEVCOAGR	224
Db	169	HKTIHRDLKPNLLDDQ--LVNKIADFGLSINIMTIDGNFLKTSGSPNYNFAPEVI-SGK	225
QY	225	ADGLAVDTGDVWAFGLVFLCVLTGNFPFEAAAGDAFAFEFVKRWQRGLPGLPQSWRR-	283
Db	226	LYAGPEVDVWSAGVILYVLMCGRLP-----FDDEFI-----PALFKKI	263
QY	284	FTPE-----ALRMFORLLALEPERRGPAKEV	309
Db	264	SNGVYTLNLYLSAGAKHLLTRMLVWPNRITTHEI	299

RESULT 10

SNFL_SCHPO STANDARD; PRT; 576 AA.

ID ID SNFL_SCHPO

AC 074536;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE SNFL-like protein kinase (EC 2.7.1.-).

GN SPCC74.03C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RC MEDLINE=21848401; PubMed=11859360;

FX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Saurous J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidaigo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

Subsequence (DEC-1996) to the EMBL/GenBank/DBJ databases.

```

[3]
RN  SEQUENCE FROM N.A.
RP  STRAIN=cv. Columbia;
RX  MEDLINE=21016720; PubMed=11130713;
RA  Salanoubat M., Lemcke K., Rieger M., Anseorge W., Unseld M.,
RA  Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA  Delisney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA  De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,
RA  Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA  Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA  Wurmbach E., Drzonek H., Efrle H., Jordan N., Bangert S.,
RA  Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA  Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA  Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordaiek G.,
RA  Reichelt J., Scharfe M., Schoen O., Barques M., Terol J., Climent J.,
RA  Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA  Cooke R., Laudie M., Berger-Lilauro C., Purnelle B., Masuy D.,
RA  de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA  Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA  Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA  Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA  Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA  Cressy T.H., Haas B., Mailer R., Wu D., Peterson J., Van Aken S.,
RA  Pai G., Millscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA  Preuss D., Lin X., Niemman W.C., Salzberg S.L., White O., Venter J.C.,
RA  Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA  Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA  Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA  Nakayama A., Yanada M., Shimpoto S., Takeuchi C., Wada T.,
RA  Watanabe A., Yanada M., Yasuda M., Tabata S.;
RT  "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT  thaliana."
RL  Nature 408:820-822 (2000).
[4]
RN  SEQUENCE OF 1-19 FROM N.A.
RP  STRAIN=cv. Columbia;
RX  MEDLINE=95115691; PubMed=7816049;
RA  le Guen L., Thomas M., Kreis M.;
RT  "Gene density and organization in a small region of the Arabidopsis
RT  thaliana genome."
RL  Mol. Gen. Genet. 245:390-396 (1994).
-!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN A SIGNAL TRANSDUCTION
CC  CASCADE REGULATING GENE EXPRESSION AND CARBOHYDRATE METABOLISM IN
CC  HIGHER PLANTS.
CC  -!- TISSUE SPECIFICITY: EXPRESSED IN ROOTS, SHOOTS AND LEAVES.
CC  -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC  SNF1 SUBFAMILY.
CC  -!- SIMILARITY: Contains 1 UBA domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M93023; AAA32736.1; -
DR  EMBL; X79707; CAA56146.1; -
DR  EMBL; AC008261; AAF26165.1; -
DR  EMBL; X94757; CAA64384.1; -
DR  FIR; JCI4446; JCI446.
DR  HSSP; Q63450; 1A06.
DR  InterPro; IPR001772; Kinase Cterm.
DR  InterPro; IPR000719; Prot Kinase.
DR  InterPro; IPR002290; Ser_Thr_kinase.
DR  InterPro; IPR001245; Tyr_kinase.
DR  InterPro; IPR000449; UBA_domain.
DR  Pfam; PF02149; KAI; 1.
DR  Pfam; PF00069; pkinase; 1.
DR  Pfam; PF00627; UBA; 1.
DR  PRINTS; PR00109; TYRKINASE.

```

```

DR  ProDom; PD000001; Prot kinase; 1.
DR  SMART; SM00220; S_TKC; 1.
DR  SMART; SM00165; UBA; 1.
DR  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR  PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR  PROSITE; PS00108; PROTEIN KINASE_ST; 1.
DR  PROSITE; PS00030; UBA; 1.
KW  Transferase; Serine/threonine-protein kinase; ATP-binding;
KW  Phosphorylation.
FT  DOMAIN 19 271 PROTEIN KINASE.
FT  DOMAIN 292 332 UBA.
FT  NP_BIND 25 33 ATP (BY SIMILARITY).
FT  BINDING 48 48 ATP (BY SIMILARITY).
FT  ACT_SITE 142 142 BY SIMILARITY.
FT  MOD_RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ  SEQUENCE 512 AA; 58373 MW; 5A186E5A0AA506DF CRC64;
Query Match 15.0%; Score 336.5; DB 1; Length 512;
Best Local Similarity 30.3%; Pred. No. 5.6e-14;
Matches 97; Conservative 59; Mismatches 111; Indels 53; Gaps 13;
QY 20 GTAPGPCAGVPLLTEDMQALTTLTAASDVTKHVELVRELKGTGKVDLVVYKGTGTM 79
DB 3 GSGTGRSGV-----ESLPNFKLRTLGISFGKVKAEHALTGKV 45
QY 80 ALKFVNKSKTK----LKNFLREYSITNSLSSPFIKVFVDVFETEDCYVFAQYAPAGD 135
DB 46 AIKILNRKIKKMEMEKVREIKILR-LFMHPHILRLVEVETPTDIYL-VMEYVNSGE 103
QY 136 LFLIIPQVGLPDTVRCVQOIGLALDFVHGQLVHRDIKPNVLLFDECHRRVKLADF 195
DB 104 LFDYIVKEKGLQDEARNFFQIISGVEYCHRNWVHRDLKPNLL-DSKC-NVKIADF 161
QY 196 GMTR--RVGCRVRSVGTIETAPETACQAGRADGLAVDTGVDVMAFGVLIFCVLTGNFPW 253
DB 162 GLSNIMEDGHFLXTSCGSPNYAAPEVI-SK---LYAGPEVDVWSCGVIIYALCGTLFP 217
QY 254 EAAGADAFEEFVRVMOGRGLPGLPSQWRRTTPALR-MFORLLALEPERGPAKEVFRF 312
DB 218 D-----DENIPNLFKKIGGIYTLPSH---LSFGARDLIPRLMVDP----- 256
QY 313 LKHELTSLSRRRSHRARKP 332
DB 257 MKRVTTPEIRQHPWFOAHLF 276

```

RESULT 5

```

CDL1_SCHPO
ID  CDL1_SCHPO STANDARD; PRT; 460 AA.
AC  Q09170; O42642;
DT  01-NOV-1995 (Rel. 32, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Protein kinase cds1 (EC 2.7.1.-) (checkpoint kinase cds1).
GN  CDL1 OR SPCL185.11C.
OS  Schizosaccharomyces pombe (Fission yeast).
OC  Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC  Schizosaccharomycetales; Schizosaccharomycetaceae;
OX  NCBI_TaxID=4896;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=95240713; PubMed=7723827;
RA  Murakami H., Okayama H.;
RT  "A kinase from fission yeast responsible for blocking mitosis in S
RT  phase."
RL  Nature 374:817-819 (1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=972;
RA  Lindsay H.D., Griffiths D.J.F., Edwards R.J., Christensen P.U.,
RA  Murray J.M., Osman F., Walworth N., Carr A.M.;
RT  "S-phase specific activation of Cds1 kinase defines a sub-pathway of

```



```

FT NP_BIND 132 140 ATP (BY SIMILARITY).
FT BINDING 155 155 ATP (BY SIMILARITY).
FT ACT_SITE 251 RPEED -> VAVPY (in isoform b).
FT VARSPLIC 402 406 /FTid=VSP_004898.
FT VARSPLIC 408 469 Missing (in isoform b).
FT VARSPLIC /FTid=VSP_004899.
SQ SEQUENCE 469 AA; 54158 MW; A71616233ABD05A9 CRC64;

Query Match
Best Local Similarity 29.3%; Score 433.5; DB 1; Length 469;
Matches 99; Conservative 79; Mismatches 141; Indels 19; Gaps 5;

QY 38 ALTERTLAASDVTHKYLRLVRELKGTGKVDLVVYKGTGTMALKFNKSKTKLKNFLRE 97
DB 111 SVTMEPIKKVLEEVYTNKOLGTGREGFKLAHQKQKORIAIKFFRPQTQQADEFRE 170
QY 98 VSIINSSSPFIKVDVVPETED--CYVPAQAYAPAGDLFDIIPP--QVGLPEDTVKR 153
DB 171 YNYSFFLSPHQNIIDTVEGFMQSSDDTAIFYEVQFCPRASREAVEATNQAGIGEANTKK 230
QY 154 CVOQLGLALDFHMGQVLRHDIKPNVLLFD-RECRVRKVLADFGMTERRVGRVSRVSGTI 212
DB 231 VFAAVLSAIFRMHNDENLVRNKLKAENILIFDANDYSKVYDFGLTRKVDTVTKYLEVYN 290
QY 213 PYTAPEVCQAGRADGLAVDTGVVWAGVLIFCVLTGNFPWEAASGADAFEEFVRWORG 272
DB 291 NYAAELCDTVNNEKLVNKSSTDIWALGIIPFYCMKGKFFWQKASIMCKPYWEWEQWLKR 350
QY 273 RLPLGSRWRRTPEPALRMFORLALPERGPAKEVREL-KHELASELRRRSHRARK 331
DB 351 KNPALPKKFPFSEKALFKKSTTPRFKORWTADMRKLAKEKLKLSVKV----- 401
QY 332 PGDRPPAAGPLRLAEPGLKRTVLTSGSGSRPAPPA 369
DB 402 ----RPEEDYVMDTASKSRQTATSSGEPQDPSAPA 435

RESULT 2
MRK4 HUMAN
ID MRK4 HUMAN STANDARD; PRT; 752 AA.
AC Q96L34; Q96JG7; Q9BYD8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE MAP/microtubule affinity-regulating kinase 4 (EC 2.7.1.27)
DE (MAP/microtubule affinity-regulating kinase like 1).
DE MARK4 OR MARKL1 OR KIAA1860.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 1).
RA Drewes G., Mandelkow E.M.;
RC TISSUE=Brain;
RT "MARK4, homologue of MARK1, MARK2 and MARK3.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=21226021; PubMed=11326310;
RA Kato T., Satoh S., Okabe H., Kitahara O., Ono K., Kihara C.,
RA Tanaka T., Tsunoda T., Yamaoka Y., Nakamura Y., Furukawa Y.;
RT "Isolation of a novel human gene, MARKL1, homologous to MARK3 and its
RT involvement in hepatocellular carcinogenesis.";
RL Neoplasia 3:4-9(2001).
RN [3]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Beghini A., Magnani I., Roversi G., Piepoli T., Diterlizzi S.,
RA Pollo B., Conti A.M.F., Cowell J.K., Finocchiaro G., Larizza L.;
RT "Neural progenitor-restricted isoform of MARKL1 gene is upregulated by

```

```

RT 19q13 amplification in human glioblastoma.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
RN [5]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Thyroid;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Makamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuhara Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96L34-1; Sequence=Displayed;
CC Name=2; Synonyms=MARKL1S;
CC IsoId=Q96L34-2; Sequence=VSP_004946;
CC -I- TISSUE SPECIFICITY: Ubiquitous. Isoform 2 is brain-specific.
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MARK
CC SUBFAMILY.
CC -I- SIMILARITY: Contains 1 kinase-associated (KAI) domain.
CC -I- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY057448; AAL23683.1; -
CC EMBL; AB049127; BAB39380.1; -
CC EMBL; AY120867; AAM55491.1; -
CC EMBL; AK075272; BAC11510.1; -
CC EMBL; AB058763; BAB47489.1; ALT_INIT.
CC HSP; Q63450; IAO6.
CC Genew; HGNC:13538; MARK4.
CC MIM; 606495; -
CC InterPro; IPR001772; Kinase_Cterm.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR000449; UBA_domain.
CC Pfam; PF02149; KAI; 1.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00627; UBA; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PSS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PSS0030; UBA; 1.
CC Transfaser; Serine/threonine-protein kinase; ATP-binding;
CC Alternative splicing.
CC DOMAIN 59 310 PROTEIN_KINASE.
CC FT DOMAIN 324 368 UBA.
CC FT DOMAIN 703 752 KAI.
CC FT NP_BIND 65 73 ATP (BY SIMILARITY).
CC FT BINDING 88 88 ATP (BY SIMILARITY).
CC FT ACT_SITE 181 181 BY SIMILARITY.
CC FT VARSPLIC 628 752 ADEPERIGGPEVTSCHLPWDQTETAPRLRFPWVSKLTSSR

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:50:56 ; Search time 17 Seconds
(without alignments)
1172.901 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPFPFRLTCCGPG.....GRADKSKGVVLATAIEICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	433.5	19.4	469	1	YWY3 CAEEL	Q11090 caenorhabdi
2	341.5	15.3	752	1	MRK4 HUMAN	O96134 homo sapien
3	337.5	15.1	619	1	SNF1 CANTR	O94168 candida tro
4	336.5	15.0	512	1	KL10 ARATH	Q38997 arabidopsi
5	335	15.0	460	1	CDS1 SCHPO	O91970 schizosacch
6	333	14.9	476	1	CHK1 MOUSE	O35280 mus musculu
7	330.5	14.8	774	1	MRK2 MOUSE	O47577 homo sapien
8	328	14.6	476	1	CHK1 HUMAN	P52497 candida alb
9	324	14.5	620	1	SNF1 CANAL	O74536 schizosacch
10	317.5	14.2	576	1	MRK3 SCHPO	P27448 homo sapien
11	317	14.2	776	1	MRK3 HUMAN	O60285 homo sapien
12	316.5	14.1	661	1	Y537 HUMAN	O91937 rattus norv
13	315	14.1	552	1	AAK2 RAT	O92265 mus musculu
14	314.5	14.0	546	1	CHK2 MOUSE	O70405 mus musculu
15	311.5	13.9	1051	1	ULK1 MOUSE	Q02723 scalce cere
16	311	13.9	502	1	RX11 SECCE	Q96555 homo sapien
17	309.5	13.8	414	1	S17A HUMAN	Q63531 rattus norv
18	309.5	13.8	735	1	AAK1 RAT	P54646 homo sapien
19	309	13.8	552	1	AAK2 HUMAN	O60782 saccharomyc
20	309	13.8	633	1	SNF1 YEAST	O96017 homo sapien
21	307	13.7	543	1	CHK2 HUMAN	O61267 drosophila
22	306.5	13.7	476	1	LOK DROME	Q00372 candida gla
23	306.5	13.7	611	1	SNF1 CANGA	Q60670 mus musculu
24	304	13.6	735	1	GA1 HUMAN	Q12236 saccharomyc
25	303.5	13.6	779	1	SNL1 MOUSE	O9rlu5 rattus norv
26	303	13.5	1081	1	KOKO YEAST	Q12263 saccharomyc
27	302	13.5	776	1	SNL1 RAT	P57059 homo sapien
28	302	13.5	1142	1	GIN4 YEAST	Q9uqb9 homo sapien
29	301.5	13.5	786	1	SNL1 HUMAN	Q9wtu3 mus musculu
30	301	13.4	309	1	STK2 HUMAN	P51812 homo sapien
31	301	13.4	733	1	GA2 MOUSE	P18654 mus musculu
32	300.5	13.4	740	1	GA3 HUMAN	
33	300.5	13.4	740	1	GA3_MOUSE	

RESULT 1

YWY3 CAEEL STANDARD; PRT; 469 AA.
AC Q11090; Q8MQF5; Q8MQF6;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative serine/threonine-protein kinase ColC4.3 in chromosome X
DE (EC 2.7.1.1-).
GN ColC4.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nelson J.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=Q11090-1; Sequence=Displayed;
CC Name=b;
CC IsoId=Q11090-2; Sequence=VSP_004898, VSP_004899;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U41025; AAM51507.1; -;
CC EMBL; U41025; AAM51508.1; -;
CC WormPep; C01C4.3a; CE30845.
CC WormPep; C01C4.3b; CE30846.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKC_1.
CC SMART; SM00219; Ty_KG_1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; FALSE NEG.
CC PROSITE; PS50011; PROTEIN KINASE DOW; 1.
CC Hypothetical protein; Transferase; Serine/threonine-protein kinase;
CC ATP-binding; Alternative splicing.
CC DOMAIN 126 397
CC PROTEIN KINASE.

34 299.5 13.4 397 1 S17A_RABIT
35 299.5 13.4 580 1 SRK1_SCHPO
36 298 13.3 856 1 UN51 CAEEL
37 296 13.2 548 1 AAK1 RAT
38 296 13.2 550 1 AAK1_HUMAN
39 294.5 13.2 1050 1 ULK1_HUMAN
40 294 13.1 733 1 KGA2_HUMAN
41 293.5 13.1 481 1 AKT2_MOUSE
42 293.5 13.1 740 1 DCK1_HUMAN
43 293.5 13.1 756 1 DCK1_MOUSE
44 292.5 13.1 481 1 AKT2_HUMAN
45 291.5 13.0 433 1 DCK1_RAT

ALIGNMENTS

Q9gm70 oryctolagus
Q94547 schizosacch
Q23023 caenorhabdi
P54645 rattus norv
Q13131 homo sapien
O75385 homo sapien
Q15349 homo sapien
Q60823 mus musculu
O15075 homo sapien
Q9jlm8 mus musculu
P31751 homo sapien
O08875 rattus norv

```
DR WPI; 2001-656860/75.
DR N-PSDB; ABL05221.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 10146; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 577 AA;
Query Match 20.2%; Score 452.5; DB 22; Length 577;
Best Local Similarity 37.4%; Pred. No. 3.9e-30;
Matches 99; Conservative 48; Mismatches 113; Indels 5; Gaps 2;
QY 53 YELVRELKGYGVKVDLVVYGTGKTKALKFVNKSKTKLNFLREVSITNLSLSPFIK 112
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 26 FDIQIVGEGWFGKILLIVHRSQSOTEMVLKAVPKPYVTLRDFRFFHYGLHGVHRHIVT 85
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 113 VFDVVFETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVOQLGLALDFMGRQLVH 172
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 86 TYDVAFTAGYVITQEVAPLGLDTSNVTDG-GGVEVYSKRVAKQLASAIIDYHSDKDIVH 144
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 RDIKPENVLFDRECRVKLADFGMTRRVRCRVKVSCTIIPYTAPEVCQAGRADGLAVDT 232
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 145 RDIKLDNLVIRSDFORIKLDFGESPTGTSTVERRNEWLPYSPEVLEIKPEGSYKADP 204
QY 233 GVDYWAEGVLIFCVLTGNFPWEAASGADAFEEFVRWQSG-----RLPGLPSQWRRTPEA 288
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 SHDYWQGVIVFVCLTGLCPWQKAASDDPRVRYLAWQGLMMPLRRTTRPLRLKLTSA 264
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 289 LRMFORLLALEPERRGPAKEVEREL 313
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 QRMFKRPFARISNRPKSLADVTKFL 289
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 15
ID ABB65543
XX ABB65543 standard; Protein; 456 AA.
XX
AC ABB65543;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 23421.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
```

```
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL09646.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 23421; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 456 AA;
Query Match 19.6%; Score 438; DB 22; Length 456;
Best Local Similarity 34.1%; Pred. No. 5.1e-29;
Matches 113; Conservative 54; Mismatches 132; Indels 32; Gaps 8;
QY 5 CPPEPPRSITCCGPTAGPGA-----GVLITEDMQALTTLTAAAD 48
Db 44 CSTPLTPTS-TSTGRLEVPGAASVSRSSYKKPKDNDGGQIHLIPDVE-LPLMTFA-- 98
QY 49 VTKHYELVRELHGTGKVDLVVYGTGKTKALKFVNKSKTKLNFLREVSITNLSLSP 108
Db 99 --DOYNIEKTLAAGCFKILLCHRPNTLVVLKANVHAELTTKFKQRFHYNELSHH 156
QY 109 FIHKVDFVFEEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMGR 168
Db 157 HILSAVAVAFQTDYVVFAMEHAPYGLASNIGPN-GLHENACKLISEQLSSALGFHSHK 215
QY 169 QLVHRDIKPENVLFDRECRVKLADFGMTRRVRCRVKVSCTIIPYTAPEVCQAGR 224
Db 216 NLVHRDLKIENILVFTPDFTRVKLCDFGATTKKGLLVHKVHTWTSVCPPEQLELIKNER 275
QY 225 ADGLAVDTGVDVNAFVGLIFCVLTGNFPWEAASGA-DAPFEFVRWQGRPLPGLPSOWRR 283
Db 276 FQCLPVS---DSWQFGILLNYLTGNFPWQSDAWVKDQSYANFMKYEQRKTKVPDNFR 332
QY 284 FTBPALRMFORLLALEPERRGPAKEVERFLK 314
Db 333 FSPRLMRCKRYLSDHPEDRCKITEVAKYMK 363
Search completed: December 5, 2003, 09:56:21
Job time : 48 secs
```

XX AC AAE32027;
 XX DT 24-MAR-2003 (first entry)
 XX DE Human kinase and phosphatase (KPP)-8.
 XX KW Human; kinase and phosphatase; KPP; cardiovascular; hypertension;
 KW myocardial infarction; angina pectoris; Alzheimer's disease; epilepsy;
 KW acquired immune deficiency syndrome; AIDS; Grave's disease; diabetes;
 KW neurological; Parkinson's disease; cirrhosis; psoriasis; gene therapy;
 KW hypercholesterolaemia; anticonvulsant; hepatotropic; lipid myopathy;
 KW cell proliferative disorder; cancer; cardiac; neuroprotective; enzyme;
 KW neurotropic; ophthalmological; anorectic; cycostatic; cataract; obesity.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FH Peptide 9..39
 XX FT /label= Signal_peptide
 XX FT Protein 40..414
 XX FT /note= "Human mature KPP protein"
 XX FT Peptide 9..38
 XX FT /label= Signal_peptide
 XX FT Protein 39..414
 XX FT /note= "Human mature KPP protein"
 XX FT Domain 98..349
 XX FT /note= "Protein kinase domain"
 XX FT Domain 271..298
 XX FT /note= "Transmembrane domain"
 XX PN WO200283709-A2.
 XX PD 24-OCT-2002.
 XX PF 05-APR-2002; 2002WO-US10818.
 XX PR 06-APR-2001; 2001US-282119P.
 XX PR 13-APR-2001; 2001US-283589P.
 XX PR 13-APR-2001; 2001US-283759P.
 XX PR 20-APR-2001; 2001US-285589P.
 XX PR 27-APR-2001; 2001US-287036P.
 XX PR 04-MAY-2001; 2001US-288608P.
 XX PR 04-MAY-2001; 2001US-288712P.
 XX PR 09-MAY-2001; 2001US-289909P.
 XX PR 17-MAY-2001; 2001US-292246P.
 XX PA (INCY-) INCYTE GENOMICS INC.
 XX PI Recipon SA, Burrill JD, Marcus GA, Zingler KA, Tang YT;
 PI Thornton M, Borowsky ML, Baughn MR, Burford N, Lee SY, Bandman O;
 PI Hafalia AJA, Yao MG, Rankumar J, Walla NK, Lu DAM, Arvizu CS;
 PI Ieon CH, Ding L, Lu Y, Gururajan R, Walsh RT, Gandhi AR;
 PI Swarnakar A, Forsythe IJ, Yue H, Au-young JK, Elliott VS, Lee S;
 XX WPI: 2003-092995/08.
 XX N-PSDB: AAD49449.
 XX PT New human kinases and phosphatases (KPP), useful for diagnosing,
 PT treating and preventing diseases or conditions associated with the
 PT aberrant KPP expression, e.g. hypertension, cancer, AIDS, diabetes,
 PT cataract, psoriasis, obesity -
 XX Claim 1; Page 162-163; 195pp; English.
 XX The invention relates to human kinases and phosphatases (KPP), and their
 CC corresponding nucleic acids. The polypeptides and polynucleotides of the
 CC invention are useful in diagnosing, treating and preventing diseases or
 CC conditions associated with the decreased expression or overexpression of
 CC KPP, such as cardiovascular (e.g. hypertension, myocardial infarction,
 CC angina pectoris), immune (e.g. acquired immune deficiency syndrome
 CC (AIDS), Grave's disease, diabetes), neurological (e.g. Parkinson's

CC disease, Alzheimer's disease, epilepsy), disorders affecting growth
 CC and development (e.g. cirrhosis, psoriasis, cataract), lipid (e.g.
 CC hypercholesterolaemia, obesity, lipid myopathies), cell proliferative
 CC disorders, or cancer. They are also useful in assessing the effects of
 CC exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of KPP. The KPP or its fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide.
 CC The microarray is useful in monitoring or measuring protein-protein
 CC interactions, drug-target interactions, and gene expression profiles.
 CC KPP DNA is used in gene therapy. The present sequence is human KPP
 XX protein.
 XX SQ Sequence 414 AA;
 Query Match 22.7%; Score 508.5; DB 24; Length 414;
 Best Local Similarity 37.1%; Pred. No. 4.1e-35;
 Matches 119; Conservative 38; Mismatches 137; Indels 27; Gaps 4;
 QY 19 PGTAPGFGAG-----VPLLTED-----MOALTTLTLAASDVTKHYE 54
 DB 40 PGLLPQTGAGGASVAVTPNLSRTQKVAVRVEDTATALQRLVELTTSRVTPVSLRDOYH 99
 QY 55 LVRELKGTGYKVDLVVYKGTGYMALEFVNKSKTKLKNFLRVSTNSUSSPFIKVF 114
 DB 100 LIRKLGSGSYGRVLLAQPHQGPAVALKLLRRDLVLRSTFLREFCVRCSAHPGLIQL 159
 QY 115 DVVFETEDCVFAQEAQAPAGDLFIIPQVGLPEDTVKCVQQLGLALDPMHGRQLVHRD 174
 DB 160 AGPLQTPRFAPFAQEAQAPCGDLGMLQER-GLPELLVKKRVVAQALAGALDFLHSGRLVHAD 218
 QY 175 IKPENVLLEDFRECKRVKLADFGMTREVGRVKRVSGTIPYTAPEVQCAQAGADGLAVDTGV 234
 DB 219 VKEDNVLVFEDPCSRVALGDLGLTRPEGSGTPAPPVPLPTAPPELCULLPPDLPKPAV 278
 QY 235 DVWAFGLVLCVLTGNFPWEASGADAFEEFVRW--QRGLPGLSQWRFTPEALRMF 292
 DB 279 DSMGLGLVLFCAATACFPMDVALAPNPEFAFAGVTTKPPQPPPPPPMDQFAPPALALL 338
 QY 293 QRLALEPERRRGPAKEVRFEL 313
 DB 339 QGLLDLPETRSPLAVLDLFL 359
 RESULT 14
 ABB61118
 ID ABB61118 standard; Protein; 577 AA.
 XX AC ABB61118;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 10146.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX	31-MAY-2001.
PD	
XX	
PF	22-NOV-2000; 2000WO-US32085.
XX	
PR	24-NOV-1999; 99US-0167482.
XX	
PA	(SUGB-) SUGEN INC.
XX	
PI	Plozman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI	Flanagan P, Clary D;
XX	
DR	WPI; 2001-343950/36.
DR	N-PSDB; AAS06707.
XX	
PT	Nucleic acids encoding human kinase polypeptides, useful for preventing
PT	diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT	neuronal-associated diseases, and microbial infections -
XX	
PS	Claim 7; Figure 2; 433pp; English.
XX	
CC	AUU03501-AAU03557 represent novel human protein kinases #1-57. The
CC	novel protein kinases have been identified as members of the tyrosine
CC	or serine/threonine kinase (PTK and STRK) families. The polynucleotides
CC	encoding protein kinases and the polypeptides may be used in the
CC	prevention, diagnosis and treatment of diseases associated with
CC	inappropriate kinase expression. For example, they may be used to treat
CC	cancers (especially cancers of haematopoietic origin), cardiovascular
CC	disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC	immune related diseases (e.g. rheumatoid arthritis), neurological
CC	disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC	Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC	disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC	Additionally, polynucleotides encoding protein kinases may be
CC	used for gene therapy and as DNA probes in diagnostic assays.
CC	The protein kinase polypeptides may be used as antigens in the production
CC	of antibodies against the protein kinases and in assays to identify
CC	modulators of protein kinase expression and activity.
XX	
SQ	Sequence 373 AA;
	Query Match 22.7%; Score 508.5; DB 22; Length 373;
	Best Local Similarity 37.1%; Pred.No. 3.6e-35;
	Matches 119; Conservative 38; Mismatches 137; Indels 27; Gaps 4;
QY	19 PGTAPEPGAG-----VFLLTSD-----MCALTRLTLAASDVTKHVE 54
Db	10 PGLLFTGTGAGSVATFNNLSRTOKVARVEDTATALQRVLVLTSTRVTRSLRDQVH 69
QY	55 LVRELGKHYGVKDVLVVYKGTGTMALKFNVNKS TKLKNFLRVSITNSLSSPFIKVF 114
Db	70 LIRKLGSYGVRLLAQPHQGQPAVALKLLRLDLVLRSTFLREFCVCGRCSAHPLQLTL 129
QY	115 DVVFETEDCVFAQAEVAPAGDLEDIIPQVGLPEDTVKECVQOLGLALDFMIGROLVHRD 174
Db	130 AGPLQTPRVFAQAEVAPCGDSLGMQLQER-GUPELLVKRWVAQLAGALDFLHSRGLVHAD 188
QY	175 IKPENVLLEDRCRRVKLADFGMTRRVGCRRVRKVSCTIPTYAPEVQCAGRADGLAVDTGV 234
Db	189 VKPNDNLVFPVCSRALVDGLGTRPEGSSPTAPPVPPLTAPPELCILLPPDTLPLRPV 248
QY	235 DWVAFGLIFCVLTGNFPWEASGADAFTEEFVRW--QRGRILGPLSQWRRTTEPALRMF 292
Db	249 DSWGVLGLVLFCAATCFPMDDVALANPEEFAFGWTTTKPQPQQPPDPWFAPPALL 308
QY	293 QRLLALEPERRRGPKEVRFEL 313
Db	309 QGLLDLDPTRSPPLAVLDLFL 329
RESULT 13	
AAE332027	
ID	AAE332027 standard. Protein: 414 AA

```
Query Match      24.4%; Score 546.5; DB 23; Length 348;
Best Local Similarity 42.9%; Pred.No. 1.8e-38;
Matches 117; Conservative 42; Mismatches 111; Indels 3; Gaps 3;
```


CC of the activity of the kinase. The proteins, nucleic acids,
 CC antibodies and modulators are useful for diagnosing, treating or
 CC monitoring cancer or aberrant cellular proliferation and/or
 CC differentiation such as lung, ovarian or brain cancer, pain (e.g.
 CC inflammation or infection) or metabolic disorder (e.g. obesity, anorexia
 CC nervosa, cachexia, lipid disorders and diabetes) or brain disorder
 CC (e.g. Alzheimer's disease, Pick's disease, Parkinsonism, Huntington's
 CC disease), neural tube defects, cerebrovascular diseases (e.g. hypoxia,
 CC ischaemia, infarction), AIDS (acquired immunodeficiency syndrome) -
 CC associated myopathy, multiple sclerosis, disorders of the heart
 CC (e.g. heart failure, angina pectoris, myocardial infarction),
 CC blood vessel disorders (e.g. arteriovenous fistula, vasculitis,
 CC syphilitic aneurysms), disorders of the bone (e.g. osteoporosis,
 CC cirrhosis, hypoparathyroidism, hyperparathyroidism, tropical sprue,
 CC idiopathic hyperglycaemia, hematopoietic disorders such as autoimmune
 CC disease including rheumatoid arthritis, osteoarthritis, multiple
 CC sclerosis, systemic lupus erythematosus, psoriasis, conjunctivitis,
 CC ulcerative colitis, asthma, allergic asthma, autoimmune uveitis,
 CC aplastic anaemia, Grave's disease, graft-versus-host disease; liver
 CC disorders e.g. Gaucher's disease, Wilson's disease; and platelet
 CC disorders. Numerous other diseases that can be treated or prevented
 CC are listed in the specification. The present sequence is the protein
 CC kinase 32374.
 XX
 SQ Sequence 346 AA;

Query Match 81.4%; Score 1822; DB 23; Length 346;
 Best Local Similarity 100.0%; Pred. No. 8.9e-148;
 Matches 346; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 79 MALKFVNKSTKLNKFNREVSITNSLSSSPFIKVFDFWETEDCYVFAOEYAPAGDLFD 138
 Db 1 MALKFVNKSTKLNKFNREVSITNSLSSSPFIKVFDFWETEDCYVFAOEYAPAGDLFD 60
 QY 139 IIPQVGLPDTVKRCVQQLGLALDFMHGRLVHRDIKPNVLLFDRCEKRVKLADFGMT 198
 Db 61 IIPQVGLPDTVKRCVQQLGLALDFMHGRLVHRDIKPNVLLFDRCEKRVKLADFGMT 120
 QY 199 RRVGCRVKRVSIGTIPYTAPEVCQAGRADGLAVDTGVDFWAFGLIFCVLTGNFPWEAASG 258
 Db 121 RRVGCRVKRVSIGTIPYTAPEVCQAGRADGLAVDTGVDFWAFGLIFCVLTGNFPWEAASG 180
 QY 259 ADAFFEEFVWQGRGLPLGSQWRRTPEPALRMFORLLALEPERRGPAKEVFRFLKHELT 318
 Db 181 ADAFFEEFVWQGRGLPLGSQWRRTPEPALRMFORLLALEPERRGPAKEVFRFLKHELT 240
 QY 319 SELRRRPSHRAKPPGDRPPAAGPLRLLEAPGLKRTVLTSSGSGSRPAPPAVGSVPLPVP 378
 Db 241 SELRRRPSHRAKPPGDRPPAAGPLRLLEAPGLKRTVLTSSGSGSRPAPPAVGSVPLPVP 300
 QY 379 VPVPVPVPVPVPPEGLAPQPGRTDGRADKSGQVVLATAIEICV 424
 Db 301 VPVPVPVPVPPEGLAPQPGRTDGRADKSGQVVLATAIEICV 346

RESULT 9
 AAU03506
 ID AAU03506 standard; Protein; 572 AA.
 XX
 AC AAU03506;
 XX
 DT 12-SEP-2001 (first entry)
 XX
 DE Human protein kinase #6.
 XX
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200138503-A2.

XX PD 31-MAY-2001.
 XX PF 22-NOV-2000; 2000WO-US32085.
 XX PR 24-NOV-1999; 99US-0167482.
 XX PA (SUGP-) SUGEN INC.
 XX PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 DR WPI: 2001-343950/36.
 DR N-PSDB; AAS06706.
 XX
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 XX
 PS Claim 7; Figure 2; 433pp; English.
 XX
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of hematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 XX
 SQ Sequence 572 AA;

Query Match 24.5%; Score 547.5; DB 22; Length 572;
 Best Local Similarity 37.6%; Pred. No. 2.8e-38;
 Matches 130; Conservative 49; Mismatches 134; Indels 33; Gaps 7;
 QY 34 EDMQALTRLTAASDVTKHYELVRELKGTGTYKVDLVVYKGTCTKMLKPVNKSMTKLKN 93
 Db 139 EDMMTLSAQTLVRAEVDLEYEVRPLGQGRYGRVLLVTHRQKGTPLALKQLPKPRTSLRG 198
 QY 94 FLREVSITNSLSSSPFIKVFDFWETEDCYVFAOEYAPAGDLFDIIPQVGLPDTVKR 153
 Db 199 FLYEFCVGLSLGASHAIVTAIGIGIESAHSYSELTPEVLHGDDLMATIQPKVGUPQPAVHR 258
 QY 154 CVQQLGLALDFMHGRLVHRDIKPNVLLFDRCEKRVKLADFGMTRRVGRVKRVSGTIP 213
 Db 259 CAAQLASALEYIHARGLVYRDLKPNVLCDCPACRFRKLTDFGHTPRGTLRLAGPPIP 318
 QY 214 YTAPEVC-QAGRADGLAVDTGVDFWAFGLIFCVLTGNFPWE--AAGAAFAEEFVWQVR 271
 Db 319 YTAPELCAPPPLPEGLPIQPALDAWALGVLLFCLLTGYFPWDRPLAEDPFYEDFLIWA 378
 QY 272 GRLP-GLPSQWRRTPEPALRMFORLLALEPERRGPAKEVFRFLKHELTSELRRRPSHRA 330
 Db 379 SGQFRDRPQFWGLAPADALLRGLDDEPHRRSAVIAI----REHLGRFWQREG---- 430
 QY 331 KPFGDRPPAAGPLRLLEAPGLKRTVLTSSGSGSRPAPPAVGSVPLP 376
 Db 431 -----EAEAVGAVEE---EAGQA-----PAVATEVLP 454

RESULT 10
 ABG30935
 ID ABG30935 standard; Protein; 348 AA.

```

Db      181  LLDRECRVRKADFGMTRRVGCRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG 240
QY      241  VLIFCVLTGNFPPWEAASGADAFEEFVRWQGRGLPGLPSQWRRTFTEPALRMFORLLALEP 300
Db      241  VLIFCVLTGNFPPWEAASGADAFEEFVRWQGRGLPGLPSQWRRTFTEPALRMFORLLALEP 300
QY      301  ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLBAPGLKRTVLTESG 360
Db      301  ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLBAPGLKRTVLTESG 360
QY      361  SGRSPAPPVSGVPLPVVPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 420
Db      361  SGRSPAPPVSGVPLPVVPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 420
QY      421  EICV 424
Db      421  EICV 424

RESULT 7
AAU83904
ID      AAY83904 standard; Protein; 417 AA.
AC      AAY83904;
XX
DT      05-JUL-2000 (first entry)
XX
DE      Rat PKS protein.
XX
KW      PKS protein; Wistar rat; cerebral nervous system disease;
KW      neurological function-related disease.
XX
OS      Rattus sp.
XX
PN      JP2000060571-A.
XX
PD      29-FEB-2000.
XX
FF      20-AUG-1998; 98JP-0249064.
XX
PR      20-AUG-1998; 98JP-0249064.
XX
PA      (MITU ) MITSUBISHI CHEM CORP.
XX
DR      WPI; 2000-249682/22.
DR      N-PSDB; AAA09825.
XX
PT      Novel mammalian peptide and a polynucleotide encoding it - useful for
PT      treatment and diagnosis of cerebral nervous system diseases and
PT      neurological function-related diseases
XX
PS      Claim 1; Page 10-11; 15pp; Japanese.
XX
CC      This sequence represents the PKS protein from Wistar rats. The peptide
CC      is useful for treatment and diagnosis of cerebral nervous system diseases
CC      and neurological function-related diseases.
XX
SQ      Sequence 417 AA;

Query Match          93.1%; Score 2084.5; DB 21; Length 417;
Best Local Similarity 94.6%; Pred. No. 3.6e-170;
Matches 401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;

QY      1  MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
Db      1  MSVGCPEPEPLHSLPCCGPGAAPVPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
QY      61  KGTGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSPFFIKVDFVVFET 120
Db      61  KGTGKVDLVAYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSPFFIKVDFVVFET 120
QY      121  EDCYVFAQYAPAGDLFDIIPPQVGLPDTVKRCVQQLGLALDFMHSRQLVHRDIKPENV 180

```

```

Db      121  EDCYVFAQYAPAGDLFDIIPPQVGLPDTVKRCVQQLGLALDFMHSRQLVHRDIKPENV 180
QY      181  LLDRECRVRKADFGMTRRVGCRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG 240
Db      181  LLDRECRVRKADFGMTRRVGCRVKRVSGTIPYTAPEVCQAGRADGLAVDTGVDVWAFG 240
QY      241  VLIFCVLTGNFPPWEAASGADAFEEFVRWQGRGLPGLPSQWRRTFTEPALRMFORLLALEP 300
Db      241  VLIFCVLTGNFPPWEAASGADAFEEFVRWQGRGLPGLPSQWRRTFTEPALRMFORLLALEP 300
QY      301  ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLBAPGLKRTVLTESG 360
Db      301  ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLBAPGLKRTVLTESG 360
QY      361  SGRSPAPPVSGVPLPVVPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 420
Db      361  SGRSPAPPVSGVPLPVVPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPVPPV 420
QY      421  EICV 424
Db      414  EICV 417

RESULT 8
AAU76323
ID      AAU76323 standard; Protein; 346 AA.
AC      AAU76323;
XX
DT      21-MAY-2002 (first entry)
XX
DE      Human protein kinase 32374.
XX
KW      Human; protein kinase; 32374; 18431; cancer; infection;
KW      cellular proliferation; pain; inflammation; metabolic disorder; obesity;
KW      lipid disorder; diabetes; brain disorder; Alzheimer's disease;
KW      neural tube defect; cerebrovascular disease; hypoxia; AIDS;
KW      acquired immunodeficiency syndrome associated myopathy; heart disorder;
KW      multiple sclerosis; angina pectoris; blood vessel disorder; vasculitis;
KW      bone disorder; osteoporosis; haematopoietic disorder; autoimmune disease;
KW      rheumatoid arthritis; osteoarthritis; multiple sclerosis;
KW      systemic lupus erythematosus; asthma; Grave's disease; liver disorder;
KW      graft-versus-host disease.
XX
OS      Homo sapiens.
XX
PN      WO200210401-A2.
XX
PD      07-FEB-2002.
XX
PF      27-JUL-2001; 2001WO-US23653.
XX
PR      28-JUL-2000; 2000US-221543P.
XX
PA      (MILL-) MILLENNIUM PHARM INC.
XX
PI      Meyers R, Kapeller-Libermann R, Silos-Santiago I;
XX
DR      WPI; 2002-188748/24.
DR      N-PSDB; ABK10101.
XX
PT      Novel protein kinase nucleic acid molecules and the encoded proteins
PT      for diagnosing and treating cellular proliferative, bone, immune,
PT      cardiovascular, liver, pain or metabolic disorders and identifying
PT      modulators -
XX
PS      Claim 4; Fig 1; 141pp; English.
XX
CC      The invention relates to an isolated protein kinase polypeptide termed as
CC      32374 or 18431, their allelic variants, fragments and nucleic acids
CC      encoding them. Also included are a host cell containing the nucleic
CC      acid, an antibody that binds to the protein kinase and modulators

```

protein (STPKP) and its nucleic acid. The STPKP sequences are useful for identifying test compounds, that may act as agonists or antagonists at the receptor site and which can be regulated to provide therapeutic effects. The reagent (e.g. antibody or antisense oligonucleotide, which decreases the expression of human STPKP gene or decreases the levels of STPKP protein), STPKP agonist, STPKP protein or expression vector is useful for treating a patient with a chronic obstructive pulmonary disease (COPD), central nervous system (CNS) disorder, cancer, metabolic disease, diabetes, obesity or urology disorders. The vectors comprising STPKP DNA and reagents are useful for preparing a medicament for modulating the activity of STPKP in the diseases. These diseases include multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy, urinary incontinence, carcinoma, leukaemia, or benign prostatic hypoplasia. STPKP gene is used in gene therapy. The present sequence is human STPKP protein.

XX Sequence 424 AA;
 Query Match 99.8%; Score 2235; DB 24; Length 424;
 Best Local Similarity 99.8%; Pred. No. 4.6e-183;
 Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTKHVELVRELG 60
 DB 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTKHVELVRELG 60
 QY 61 KGTYGKVDLVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSTSPFFIIKVDVVFET 120
 DB 61 KGTYGKVDLVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSTSPFFIIKVDVVFET 120
 QY 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 DB 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 QY 181 LLFDRECRVRLADFGMTRRVGCRVKSITPIYTAPEVCOAGRADGLAVDTGVDVWAFG 240
 DB 181 LLFDRECRVRLADFGMTRRVGCRVKSITPIYTAPEVCOAGRADGLAVDTGVDVWAFG 240
 QY 241 VLIFCVLTGNFPWEAASGADAFFEFVWQGRGLPLPSQWRRTTEPALRMFQLLALEP 300
 DB 241 VLIFCVLTGNFPWEAASGADAFFEFVWQGRGLPLPSQWRRTTEPALRMFQLLALEP 300
 QY 301 ERRGPAKEVFRFLKHELTSELRRPSPHRRKPPGDRPPAAGPLLEAPGLKRVLTESG 360
 DB 301 ERRGPAKEVFRFLKHELTSELRRPSPHRRKPPGDRPPAAGPLLEAPGLKRVLTESG 360
 QY 361 SGRSPAPAVGSVPLVPVPVPVPVPVPPGLAPQGPGRDGRADKSGQVVLATAI 420
 DB 361 SGRSPAPAVGSVPLVPVPVPVPVPVPPGLAPQGPGRDGRADKSGQVVLATAI 420
 QY 421 EICV 424
 DB 421 EICV 424

RESULT 6

AAE32010

ID AAE32010 standard; Protein; 424 AA.

XX AAE32010;

AC AAE32010;

DT 24-MAR-2003 (first entry)

DE Human serine/threonine protein kinase-like protein (STPKP) #1.

XX Human; serine/threonine protein kinase-like protein; STPKP; diabetes;
 KW central nervous system; CNS; metabolic disease; urology disorder; COPD;
 KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
 KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
 KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hypoplasia;
 KW anorectic; neuroprotective; nootropic; cytostatic; uteropathic.
 XX Homo sapiens.

XX Key Location/Qualifiers
 PH Domain 53..309
 FT /note= "Eukaryotic protein kinase domain"
 FT Region 59..82
 FT /note= "Protein-kinase-ATP region"
 FT Binding-site 82
 FT /note= "ATP binding site"
 FT Region 170..180
 FT /note= "Protein-kinase-ST region"
 FT Active-site 174
 XX WO200283982-A2.
 XX 24-OCT-2002.
 XX 12-APR-2002; 2002WO-EP04080.
 XX 12-APR-2001; 2001US-283189P.
 PR 04-JUN-2001; 2001US-294989P.
 PR 30-JUN-2001; 2001US-308094P.
 PR 08-FEB-2002; 2002US-354574P.
 XX (FARB) BAYER AG.
 XX Koehler RH;
 XX WPI; 2003-093018/08.
 DR N-PSDB; AAD49416.
 XX New polynucleotides encoding serine/threonine protein kinase-like protein (STPKP) and its encoded protein, useful for identifying modulators of STPKP activity, and in gene therapy for treating e.g. cancer, obesity or diabetes
 Claim 18; Page 147-149; 161pp; English.
 The invention relates to human serine/threonine protein kinase-like protein (STPKP) and its nucleic acid. The STPKP sequences are useful for identifying test compounds, that may act as agonists or antagonists at the receptor site and which can be regulated to provide therapeutic effects. The reagent (e.g. antibody or antisense oligonucleotide, which decreases the expression of human STPKP gene or decreases the levels of STPKP protein), STPKP agonist, STPKP protein or expression vector is useful for treating a patient with a chronic obstructive pulmonary disease (COPD), central nervous system (CNS) disorder, cancer, metabolic disease, diabetes, obesity or urology disorders. The vectors comprising STPKP DNA and reagents are useful for preparing a medicament for modulating the activity of STPKP in the diseases. These diseases include multiple sclerosis, Alzheimer's disease, Parkinson's disease, epilepsy, urinary incontinence, carcinoma, leukaemia, or benign prostatic hypoplasia. STPKP gene is used in gene therapy. The present sequence is human STPKP protein.

SQ Sequence 424 AA;

Query Match 98.6%; Score 2207; DB 24; Length 424;

Best Local Similarity 98.6%; Pred. No. 1.2e-180;

Matches 418; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTKHVELVRELG 60
 DB 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTKHVELVRELG 60
 QY 61 KGTYGKVDLVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSTSPFFIIKVDVVFET 120
 DB 61 KGTYGKVDLVYKGTGTMALKFNKSKTKLKNFLREVSITNSLSSTSPFFIIKVDVVFET 120
 QY 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 DB 121 EDCVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDPMHGRQLVHRDIKPEV 180
 QY 181 LLFDRECRVRLADFGMTRRVGCRVKSITPIYTAPEVCOAGRADGLAVDTGVDVWAFG 240

FT Domain 53..309 /note= "Eukaryotic protein kinase domain"
 FT Domain 54..303
 FT Domain /note= "Protein kinase domain"
 FT Domain 55..304 /note= "Protein kinase domain"
 FT Region 128..141
 FT Region /note= "Tyrosine kinase catalytic site"
 FT Region 164..182
 FT Active-site /note= "Tyrosine kinase catalytic site"
 FT 170..182
 FT Domain /note= "Protein kinase Ser/Thr active site domain"
 FT 228..248
 FT Region /note= "Transmembrane domain"
 FT 234..256
 FT /note= "Tyrosine kinase catalytic site"
 XX WO200218557-A2.
 XX 07-MAR-2002.
 XX 31-AUG-2001; 2001WO-US27219.
 XX 31-AUG-2000; 2000US-229873P.
 XX 08-SEP-2000; 2000US-231357P.
 XX 14-SEP-2000; 2000US-232654P.
 XX 22-SEP-2000; 2000US-234902P.
 XX 23-SEP-2000; 2000US-236499P.
 XX 06-OCT-2000; 2000US-238389P.
 XX 13-OCT-2000; 2000US-240542P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Bandman O, Nguyen DB, Walia NK, Hafalia AJA, Yao MG, Gandhi AR;
 PI Gururajan R, Ding LS, Patterson C, Yue H, Baughn MR, Tribouley CM;
 PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YI;
 PI Azimzai Y, Burrill JD, Marcus GA, Zingler KA, Lu DAM, Lal PG;
 PI Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
 PI Burford N;
 XX WPI; 2002-329769/36.
 DR N-PSDB; AAD34298.
 XX New human kinases, useful for diagnosing, treating or preventing immune
 PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as
 PT leukemia or lymphoma)
 XX Claim 56; Page 141-142; 218pp; English.
 XX The present invention relates to human kinases (PKIN) and polynucleotides
 CC encoding such proteins. PKIN sequences of the invention are useful for
 CC diagnosing, treating or preventing disorders associated with aberrant
 CC expression of PKIN, particularly immune system disorders (e.g. acquired
 CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
 CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
 CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, sarcoma),
 CC and developmental disorders (e.g. Down's syndrome). They are also used
 CC in gene therapy and protein therapy. The present sequence is human
 CC PKIN-1 protein.
 XX Sequence 424 AA;
 SQ
 Query Match 99.8%; Score 2235; DB 23; Length 424;
 Best Local Similarity 99.8%; Pred. No. 4.6e-183;
 Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSVCGPEPEPRSRITCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTWKHYELVRELG 60
 DB 1 MSVCGPEPEPRSRITCCGPGTAPGAGVPLLTEDMQALTTLTAAADVTWKHYELVRELG 60
 QY 61 KGTYGKVDLVVYKGTGTMALFVNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET 120

Db 61 KGTYGKVDLVVYKGTGTMALFVNKSKTKLKNFLREVSITNSLSPPFIKVFDDVVFET 120
 QY 121 EDCYVFAQBYAPAGDLFDIIPQVGLPEDITVKRCVQQLGLALDFMHRGQLVHRDIKPENV 180
 Db 121 EDCYVFAQBYAPAGDLFDIIPQVGLPEDITVKRCVQQLGLALDFMHRGQLVHRDIKPENV 180
 QY 181 LLDRECRVRKVLADFGMTTRVRCRVKRVSGTIPYTAPEVCAQAGRADGLAVDTGVDVWAFG 240
 Db 181 LLDRECRVRKVLADFGMTTRVRCRVKRVSGTIPYTAPEVCAQAGRADGLAVDTGVDVWAFG 240
 QY 241 VLIFCVLTGNFWEAASGADAFEEFVRVQGRGLPGLPSQWRRTFEPALRMFORLLALEP 300
 Db 241 VLIFCVLTGNFWEAASGADAFEEFVRVQGRGLPGLPSQWRRTFEPALRMFORLLALEP 300
 QY 301 ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLLEAPGPKRTVLTESG 360
 Db 301 ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLLEAPGPKRTVLTESG 360
 QY 361 SGSRPAPPAGVSVPLPVVPVVPVPEPLAPQGPGRDGRADKSKGOVLATAI 420
 Db 361 SGSRPAPPAGVSVPLPVVPVVPVPEPLAPQGPGRDGRADKSKGOVLATAI 420
 QY 421 EICV 424
 Db 421 EICV 424
 RESULT 5
 AAE32011
 ID AAE32011 standard; Protein; 424 AA.
 XX AAE32011;
 AC AAE32011;
 XX 24-MAR-2003 (first entry)
 DT Human serine/threonine protein kinase-like protein (STPKP) #2.
 DE Human serine/threonine protein kinase-like protein; STPKP; diabetes;
 XX central nervous system; CNS; metabolic disease; urology disorder; COPD;
 KW chronic obstructive pulmonary disease; multiple sclerosis; gene therapy;
 KW Alzheimer's disease; Parkinson's disease; urinary incontinence; cancer;
 KW epilepsy; obesity; carcinoma; leukaemia; benign prostatic hyperplasia;
 KW anorectic; neuroprotective; neurotropic; cytosstatic; uropathic.
 XX Homo sapiens.
 OS WO200283882-A2.
 XX 24-OCT-2002.
 PD 12-APR-2002; 2002WO-EF04080.
 XX 12-APR-2001; 2001US-283189P.
 PR 04-JUN-2001; 2001US-294989P.
 PR 30-JUL-2001; 2001US-308094P.
 PR 08-FEB-2002; 2002US-354574P.
 XX (FARB) BAYER AG.
 PA Koehler RH;
 XX WPI; 2003-093018/08.
 DR N-PSDB; AAD49419.
 XX New polynucleotides encoding serine/threonine protein kinase-like
 PT protein (STPKP) and its encoded protein, useful for identifying
 PT modulators of STPKP activity, and in gene therapy for treating e.g.
 PT cancer, obesity or diabetes
 XX Claim 18; Page 157-159; 161pp; English.
 XX The invention relates to human serine/threonine protein kinase-like

QY 61 KGYGVKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSTNSLSPPFLIKVDFVVFET 120
 DB 61 KGYGVKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSTNSLSPPFLIKVDFVVFET 120
 QY 121 EDCVFAQYAPAGDLFDIIPPOVGLPEDIYKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
 DB 121 EDCVFAQYAPAGDLFDIIPPOVGLPEDIYKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
 QY 181 LFDRECRVRKVLADFGMTTRVGRVGRVSGTIPTTAPEVQAGRADGLAVDTGVDVWAFG 240
 DB 181 LFDRECRVRKVLADFGMTTRVGRVGRVSGTIPTTAPEVQAGRADGLAVDTGVDVWAFG 240
 QY 241 VLIFCVLTGNFWEAASGADAFEEFVRWGRGLPGLPSQWRRTTEPALRMFORLLALEP 300
 DB 241 VLIFCVLTGNFWEAASGADAFEEFVRWGRGLPGLPSQWRRTTEPALRMFORLLALEP 300
 QY 301 ERGPAKEVFRFLKHELTSELRRPSSHARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
 DB 301 ERGPAKEVFRFLKHELTSELRRPSSHARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
 QY 361 GSGRPAPPVGVSVPLPVVPVPPVPPVPEPGLAPQGPGRDGRADKSGQVVLATAI 420
 DB 361 GSGRPAPPVGVSVPLPVVPVPPVPPVPEPGLAPQGPGRDGRADKSGQVVLATAI 420
 QY 421 EICV 424
 DB 421 EICV 424
 RESULT 3
 AAU03544
 ID AAU03544 standard; Protein; 424 AA.
 AC AAU03544;
 XX
 DT 12-SEP-2001 (first entry)
 DE Human protein kinase #44.
 XX
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder.
 OS Homo sapiens.
 XX
 XX WO200138503-A2.
 PN
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32085.
 XX
 PR 24-NOV-1999; 99US-0167482.
 XX
 PA (SUGEN-) SUGEN INC.
 XX
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 XX
 DR WPI; 2001-343950/36.
 DR N-PSDB; AAS06744.
 XX
 PT Nucleic acids encoding human kinase polypeptides, useful for preventing
 PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 PT neuronal-associated diseases, and microbial infections -
 XX
 PS Claim 7; Figure 2; 433pp; English.
 XX
 CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
 CC novel protein kinases have been identified as members of the tyrosine
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
 CC encoding protein kinases and the polypeptides may be used in the

CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate kinase expression. For example, they may be used to treat
 CC cancers (especially cancers of haematopoietic origin), cardiovascular
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 CC immune related diseases (e.g. rheumatoid arthritis), neurological
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 CC Additionally, polynucleotides encoding protein kinases may be
 CC used for gene therapy and as DNA probes in diagnostic assays.
 CC The protein kinase polypeptides may be used as antigens in the production
 CC of antibodies against the protein kinases and in assays to identify
 CC modulators of protein kinase expression and activity.
 XX
 SQ Sequence 424 AA;
 Query Match 99.8%; Score 2235; DB 22; Length 424;
 Best Local Similarity 99.8%; Pred. No. 4.6e-183;
 Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSVGCPEPEPPSRSLTCGPGTAPGAGVPLLTEDMQALTRLTAAASDVTKHYELVRELG 60
 DB 1 MSVGCPEPEPPSRSLTCGPGTAPGAGVPLLTEDMQALTRLTAAASDVTKHYELVRELG 60
 QY 61 KGYGVKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSTNSLSPPFLIKVDFVVFET 120
 DB 61 KGYGVKVDLVVYKGTGTMALKFVNKSKTKLKNFLREVSTNSLSPPFLIKVDFVVFET 120
 QY 121 EDCVFAQYAPAGDLFDIIPPOVGLPEDIYKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
 DB 121 EDCVFAQYAPAGDLFDIIPPOVGLPEDIYKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
 QY 181 LFDRECRVRKVLADFGMTTRVGRVGRVSGTIPTTAPEVQAGRADGLAVDTGVDVWAFG 240
 DB 181 LFDRECRVRKVLADFGMTTRVGRVGRVSGTIPTTAPEVQAGRADGLAVDTGVDVWAFG 240
 QY 241 VLIFCVLTGNFWEAASGADAFEEFVRWGRGLPGLPSQWRRTTEPALRMFORLLALEP 300
 DB 241 VLIFCVLTGNFWEAASGADAFEEFVRWGRGLPGLPSQWRRTTEPALRMFORLLALEP 300
 QY 301 ERGPAKEVFRFLKHELTSELRRPSSHARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
 DB 301 ERGPAKEVFRFLKHELTSELRRPSSHARKPPGDRPPAAGPLRLLEAPGLKRTVLTESG 360
 QY 361 GSGRPAPPVGVSVPLPVVPVPPVPPVPEPGLAPQGPGRDGRADKSGQVVLATAI 420
 DB 361 GSGRPAPPVGVSVPLPVVPVPPVPPVPEPGLAPQGPGRDGRADKSGQVVLATAI 420
 QY 421 EICV 424
 DB 421 EICV 424
 RESULT 4
 AAU03544
 ID AAU03544 standard; Protein; 424 AA.
 AC AAU03544;
 XX
 DT 16-JUL-2002 (first entry)
 DE Human PKIN-1 protein.
 XX
 KW Human; kinase; enzyme; PKIN-1 protein; immune system disorder; anaemia;
 KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
 KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
 KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
 KW leukemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
 KW Down's syndrome; gene therapy; protein therapy; cytostatic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 PH

PT markers for restriction fragment length polymorphism analysis and in
PT forensic biology -
XX Claim 1; Page 35-36; 37pp; English.
XX
XX The invention discloses an isolated human kinase polypeptide, designated
CC novel human protein (NHP), and the polynucleotide encoding it. Kinases
CC mediate phosphorylation of a wide variety of proteins and compounds in
CC the cell and are also involved in a range of regulatory pathways. The
CC polynucleotide is useful in therapeutic, diagnostic and pharmacogenomic
CC applications, and for identifying compounds that modulate, i.e. act as
CC agonists or antagonists of, the gene expression or gene product
CC activity. The polynucleotide is also useful as a probe in microarrays
CC or other cloning and/or assay formats, for screening collections of
CC genetic material from patients who have a particular medical condition,
CC for identifying mutations associated with a particular disease and also
CC as a diagnostic or prognostic assay. It can also be useful for the
CC detection of mutant human proteins, or inappropriately expressed
CC proteins, for the diagnosis of disease, for screening for drugs effective
CC in perturbing the normal function of the protein, for generation of
CC antibodies and as reagents in assays for screening for compounds that can
CC be used as pharmaceutical agents in the therapeutic treatment of mental,
CC biological or medical disorders and diseases. The sequence presented is
CC the novel human protein (NHP) which is encoded by a polynucleotide
CC isolated from human brain and skeletal muscle cDNA libraries.
XX
SQ Sequence 424 AA;

Query Match 100.0%; Score 2239; DB 23; Length 424;
Best Local Similarity 100.0%; Pred. No. 2.1e-183;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
Db 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60
QY 61 KGYGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSTNSLSGSPFIKVFVDVVFET 120
Db 61 KGYGKVDLVVYKGTGTMALKFNKSKTKLKNFLREVSTNSLSGSPFIKVFVDVVFET 120
QY 121 EDCYVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
Db 121 EDCYVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
QY 181 LFDRECRVKLADFGMTRVGRVKRVSGTIPYTAPEVQAGRADGLAVDTGVVWAFG 240
Db 181 LFDRECRVKLADFGMTRVGRVKRVSGTIPYTAPEVQAGRADGLAVDTGVVWAFG 240
QY 241 VLIFCVLTGNFPWEAASGADAFEEFVRWGRGLPGLPSQWRRTFPALRMFORLLALEP 300
Db 241 VLIFCVLTGNFPWEAASGADAFEEFVRWGRGLPGLPSQWRRTFPALRMFORLLALEP 300
QY 301 ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLAPGLKKTVLTESG 360
Db 301 ERGPAKEVFRFLKHELTSELRRRPSHRARKPPGDRPPAAGPLRLAPGLKKTVLTESG 360
QY 361 SGRSPPAPVGSVPLPVVPVVPVVPVPEPGLAQCPGPRDTGRADKSGQVVLATAI 420
Db 361 SGRSPPAPVGSVPLPVVPVVPVVPVPEPGLAQCPGPRDTGRADKSGQVVLATAI 420
QY 421 EICV 424
Db 421 EICV 424

RESULT 2

AAU10023

ID AAU10023 standard; Protein; 424 AA.

XX

AC AAU10023;

XX 08-MAY-2002 (first entry)

DT

XX

DE Human protein kinase N protein.
XX
KW Human; protein kinase N; cytostatic; neuroprotective; cancer;
KW gene therapy; antigen; antibody; neurodegenerative disease;
KW inflammation; arteriosclerosis; psoriasis; growth disorder;
KW chromosome 16; papilloma virus infection; Alzheimer's disease.
OS Homo sapiens.
PN WO2001188148-A2.
PD 22-NOV-2001.
XX
XX 17-MAY-2001; 2001WO-US15776.
XX
XX 17-MAY-2000; 2000US-205228P.
PR 12-DEC-2000; 2000US-0734032.
PR 26-MAR-2001; 2001US-0816094.
XX
XX (APPL-) APPLERA CORP.
PA
XX Wei M, Chandramouliwara I, Ye J, Ketchum KA, Di Francesco V;
PI Beasley EM;
PI
XX WPI; 2002-089857/12.
DR N-PSDB; AAS17862, AAS17863.
XX
XX Human kinase protein and polynucleotides encoding them, useful for
PT identifying modulators of kinase polypeptides and for treating,
PT preventing, and/or diagnosing neurodegenerative diseases and cancer -
XX
XX Claim 1; Figure 2; 65pp; English.
PS
XX This sequence represents a human protein kinase N of the invention.
CC The invention comprises nucleotide and protein sequences of an isolated
CC protein which is related to the PKN kinase subfamily. The protein
CC kinase N gene is located on chromosome 16. The protein may have
CC cytostatic and neuroprotective and can be used in gene therapy
CC possibly as a human kinase protein expression or activity modulator.
CC The nucleic acids and polypeptides of the invention may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. The nucleic acids (or vectors
CC containing them) and the kinase may be used to treat disorders
CC associated with decreased expression, either by rectifying mutations or
CC deletions in a patient's genome that affect the activity of the enzyme
CC by expressing inactive proteins or to supplement the patients own
CC production of kinases. Additionally, the nucleic acids may be used to
CC produce the kinase, by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The nucleic acid and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence similar nucleic acids in
CC samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides may also be used as antigens in the
CC production of antibodies against the kinase and in assays to identify
CC modulators of kinase expression and activity. The anti-kinase
CC antibodies and antagonists may also be used to down regulate expression
CC and activity. The anti-kinase antibodies may also be used as diagnostic
CC agents for detecting the presence of kinase polypeptides in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). Disorders that may
CC be prevented, diagnosed and/or treated by the above methods include, for
CC example neurodegenerative diseases, inflammation, arteriosclerosis
CC psoriasis, cancer, papilloma virus infection, Alzheimer's disease
CC and growth disorders.
XX
SQ Sequence 424 AA;

Query Match 100.0%; Score 2239; DB 23; Length 424;

Best Local Similarity 100.0%; Pred. No. 2.1e-183;

Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60

Db 1 MSVGCPEPEPRSLTCCGPGTAPGAGVPLLTEDMQALTLRTLAASDVTKHYELVRELG 60

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:50:21 ; Search time 47 Seconds
(without alignments)
1431.916 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGCPEPPRSLTCCGPG.....GRADKSKGVVLTATAIEICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseqc 19Jun03:*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	100.0	424	23 AAU79585	Novel human protei
2	2239	100.0	424	23 AAU10023	Human protein kina
3	2235	99.8	424	22 AAU03544	Human protein kina
4	2235	99.8	424	23 AAU21706	Human PKIN-1 prote
5	2235	99.8	424	23 AAU32011	Human serine/threo
6	2207	98.6	424	24 AAU32010	Human serine/threo
7	2084.5	93.1	417	21 AAU83904	Rat PKX protein.
8	1822	81.4	346	23 AAU76323	Human protein kina
9	547.5	24.5	572	22 AAU03506	Human protein kina

10	546.5	24.4	348	23	ABG3093E	Human novel protei
11	546.5	24.4	348	23	AAE21721	Human PKIN-16 prot
12	508.5	22.7	373	22	AAU03507	Human protein kina
13	508.5	22.7	414	24	AAE32027	Human kinase and p
14	452.5	20.2	577	22	ABE61118	Drosophila melanog
15	438	19.6	456	22	ABE65543	Drosophila melanog
16	383.5	17.1	250	23	ABP78796	Serine/threonine k
17	367	16.4	761	24	ABP96068	Human protein kina
18	361	16.1	702	22	ABP70124	Drosophila melanog
19	355	15.9	366	24	ABP96083	Human protein kina
20	352	15.7	520	22	AAE65629	Novel protein kina
21	351.5	15.7	686	22	AAE65621	Novel protein kina
22	350.5	15.7	825	24	ABP97380	Human DCAMKL1-like
23	350	15.6	254	23	ABP78797	Protein kinase dom
24	349.5	15.6	252	23	ABP78821	Protein kinase dom
25	347.5	15.5	724	23	ABE04431	Murine neuronal se
26	346	15.5	817	23	AAO15419	Novel human kinase
27	346	15.5	835	23	ABE82471	Human serine/threo
28	346	15.5	835	23	AAE24133	Human kinase (PKIN
29	345.5	15.4	514	22	ABE60741	Drosophila melanog
30	345.5	15.4	514	22	ABE60742	Drosophila melanog
31	345.5	15.4	660	23	ABE04432	Murine neuronal se
32	344	15.4	512	21	AAE36157	Arabidopsis thalia
33	343.5	15.3	509	21	AAE03416	Corn putative carb
34	343.5	15.3	509	24	ABE40708	Zea mays oil trait
35	343.5	15.3	509	24	ABE40809	Oryza sativa oil t
36	343	15.3	512	23	AAO17663	A thaliana AKIN11.
37	342.5	15.3	531	24	ABE40718	Triticum aestivum
38	342	15.3	512	23	AAO17664	A thaliana AKIN11
39	342	15.3	523	21	AAE03425	Wheat putative car
40	342	15.3	523	24	ABE40719	Triticum aestivum
41	341.5	15.3	688	23	ABE04434	Human neuronal ser
42	341.5	15.3	729	19	AAE37158	Human Twenty-five
43	341.5	15.3	729	22	ABE65628	Novel protein kina
44	341.5	15.3	729	24	AAE33552	Human microtubule
45	341.5	15.3	752	22	AAE11782	Human kinase (PKIN

ALIGNMENTS

RESULT 1
AAU79585
ID AAU79585 standard; Protein; 424 AA.
AC AAU79585;
XX
XX 24-SEP-2002 (first entry)
DE Novel human protein (NHP).
XX Human; Novel human protein; NHP; kinase; regulatory pathway;
XX therapeutic; diagnostic; pharmacogenomic; antigenic.
XX Homo sapiens.
XX
XX WO200246428-A2.
XX
XX 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-US48533.
XX
XX 07-DEC-2000; 2000US-251941P.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX Walke DW, Maricar M, Yu X, Friddle CJ;
XX
XX WPI; 2002-527921/56.
XX N-PSDB; ABK86975.
XX
XX Novel nucleic acid molecule encoding a human kinase, useful in
XX therapeutic, diagnostic and pharmacogenomic applications, as DNA


```

Db 525 AAGVILYILLCGPPRSPERDQDELFINI---QVGQFFELSPYDNDISDAAKDLVRNLL 581
QY 297 ALEPERGPAKVFRLKHELTSELRRSRARXP 332
Db 582 EVDPKRYTAEQLQHPWIEVGHWTGNSQKEESP 617

RESULT 13
P92958 PRELIMINARY; PRT; 512 AA.
AC P92958;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein kinase (AKIN11 protein kinase) (AT3g29160/MXE2_16).
GN AKIN11.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Bhalerao R.P.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Nakamura Y.;
RL MEDLINE=20277480; PubMed=10819329;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cdna clones.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
RA Huan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; X99279; CAA67671.1; -
DR EMBL; AB018121; BAB01933.1; -
DR EMBL; AY070468; AAL49934.1; -
DR EMBL; AY149927; AAN31081.1; -
DR HSSP; Q63450; 1A06.
DR InterPro; IPR001772; Kinase Cterm.
DR InterPro; IPR000719; Prot_Kinase.

```

```

DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PF00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 512 AA; 58689 MW; 4BF55D3EC60F87 CRC64;

Query Match 15.4%; Score 344; DB 10; Length 512;
Best Local Similarity 28.5%; Pred. No. 1.5e-19;
Matches 120; Conservative 65; Mismatches 132; Indels 104; Gaps 20;

QY 49 VTKHYELVRELKGYKGYKVDLVVYKGTGTMALKFVNKSKTK-----LKNFLREVSTNSL 104
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 16 ILPNYKLGKTLGIGSGFKVKAIEHVVTGHKVAIKILNRRKIKNMEMEKEKVRSEIKILR-L 74
QY 105 SSSPFIKVFVDVVFETEDCYFAQAYAPAGDLFDIIPQVGLPEDTVKCVQOLGLALDF 164
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 75 FMHPHIIROYEVIETSDIYV-VMEYVSGELFDVIVEKRGLOEDEARNFFQOIIISGVY 133
QY 165 MEGROLVHRDIKPNVLLFDRECRVKLADFGWTR--RVGCRVKRVSGTIPYTAPEVCOA 222
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 134 CHRNVMVHRDLAPENLLL--DSEC-NIKLADFGLSNVRMDGHLKTCGSPNYAAPEVI-S 190
QY 223 GRADGLAVDTGVDVNAFGLVIFCVITGNPFWAASGADAFFEEFVWQGRPLPLPSQW- 281
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 GK---LYAGPEVDVMSGCVILYALLCGTLFPD-----DENIPNFKKIKGGIYTLPSHLS 242
QY 282 -----RRTEPALRM---FQ-----RLLALE-PERRGPAKEV----- 309
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 SEARDLIPRLMIVDPVKRITIPRIQHEWFQTHLPYLAISPDPDTVQAKKINEEIVQEV 302
QY 310 --FRFLKHELTSELRRRP-----SHRARKPPGDRPPAAGPLRLAEPGLKRTV 355
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 303 VNMGFDNRQVLESIRNRNTQNDATVTYVLLDNRFRVPSG-----YLES--EFQET- 350
QY 356 LTESGSGSRPAPAVGSVLPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPVFPV 413
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 -TDSGN-----PMRTPEAGASPVGHWTIPAHVDHYGLGARSQ 386
QY 414 V 414
Db 387 V 387

RESULT 14
Q9ZNT4 PRELIMINARY; PRT; 508 AA.
AC Q9ZNT4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OSK3.
GN OSK3 OR OSK5. (Rice).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99086251; PubMed=9870704;
RA Takano M., Kajiya-Kanegae H., Funatsuki H., Kikuchi S.;
RT "Rice has two distinct classes of protein kinase genes related to SNF1
RT of Saccharomyces cerevisiae, which are differently regulated in early

```

RT for large proteins in vitro.;

RL DNA Res. 7:347-355(2000).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AB051552; BAB21856.1; -.

DR HSP; Q63450; 1A06.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_Thr_kinase.

DR Pfam; PF00069; pkinase.1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SMO0220; S_TK; 1.

DR PROSITE; PS0107; PROTEIN KINASE ATP; 1.

DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.

DR PROSITE; PS0108; PROTEIN KINASE ST; 1.

KW Hypothetical protein; ATP-binding; kinase;

KW Serine/threonine-protein kinase; Transferase.

FT NON TER 1

SQ SEQUENCE 608 AA; 69361 MW; 809C48963A8BD9CD CRC64;

Query Match 15.5%; Score 346; DB 4; Length 608;

Best Local Similarity 34.7%; Pred. No. 1.3e-19;

Matches 96; Conservative 48; Mismatches 115; Indels 18; Gaps 9;

QY 44 LAASDVTKHYELRELKGTGKGVLDLVYKGTGTAKLKFVNKSKTKLKNFL--REVSIT 101

DB 307 IIAANVEKHVETGRVIGDGNFAVVKECRHRETROAYAMKIIDKSLKGMVDSEILII 366

QY 102 NSLSSSFILKVFVDETDCYFAQYAPAGDLFDLIIPQVGLPDTVKRCVQQLGLA 161

DB 367 QSL-SHENIVKLHE-VYETDMEIYLLEYVQGGDLFDALIESVKFPBPDAALMIMDLCKA 424

QY 162 LDFMHGQLVHRDIKPNVLLFDECR--RVKLAQFGTRRVGCRVKVSGTIPTTAPEV 219

DB 425 LVNMDKSIIVHRDLKPNLLVORNEKSTIKLADFLGAKHVPIPTVCGTPTTYVAPEI 484

QY 220 CQAGRADGLAVDTGVNFAFGLVLCVLTGNFWEAAS-GADAFEEFVRWQGRGLFCLP 278

DB 485 L-SEKYGLE---VDMWAAGVILYLLCGFPFPRSPERQDELFNII---QLGHFEFLP 536

QY 279 SOWERFTEPALRMFORLLALEPERGPAKEVERFLKH 315

DB 537 PYWNIDSAADKLVSLVVDPKKRYAHQV---LQH 570

RESULT 11

Q8CIP4 PRELIMINARY; PRT; 752 AA.

ID Q8CIP4

AC Q8CIP4

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE MAP/microtubule affinity-regulating kinase 4L.

GN MARK4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Swiss Webster;

RA Beghini A., Moroni R.F., Larizza L.;

RT "Identification of a novel mouse gene MARK4 homologous to human MARK1.";

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY151083; AAN60072.1; -.

KW Kinase.

SQ SEQUENCE 752 AA; 82643 MW; 185FDAE3F0D627DD CRC64;

Query Match 15.4%; Score 345.5; DB 11; Length 752;

Best Local Similarity 26.0%; Pred. No. 1.9e-19;

Matches 108; Conservative 71; Mismatches 155; Indels 81; Gaps 15;

QY 52 HYELVRELKGTGKGVLDLVYKGTGTAKLKFVNKSK---TKLKNFREVSIITNSLSSP 108

DB 58 NYRLRLTIGKGNFAKVKLARHILTGREVAIKIIDKQLNPSSQLKLFREVRIMKGL-NHP 116

QY 109 FILKVDVVFETEDCYFAQYAPAGDLFDLIIPQVGLPDTVKRCVQQLGLADFMHGR 169

DB 117 NIVKLFEVI-ETKTLVLMVEYASAGEVFDYLVSHGRMKEKEARAKFRQIVSAVHYCHQK 175

QY 169 QLVHRDIKPNVLLFDECRVRLADFGMTR--VGRVKRVSGTIPTTAPEVQOAGRAD 226

DB 176 NIVHRDLKAENLLL-DAEA-NIKIADFGFNEFTLGSLDTFCGSPPYAAPELFGQKKYD 233

QY 227 GLAVDTGVNFAFGLVLCVLTGNFWEAASGADAFEEFVRWQGRGLPGLPSQWRFFTE 286

DB 234 G----PEVDINSLGVILYTLVSGSLPDGHN---LKLAEERVLRGKY-RVPFYMSTDCB 284

QY 287 PALRMFORLLALEPERGPAKEVER----- 311

DB 285 SILR--RFLVLPNAKRCITLEQIMKMKWINIGYEGELKPYTEPEDFGDKRIEVMVGM 341

QY 312 -FLKHELTSELRRPSH-----BARKPGDR-PPAAGPLRLLEAPGLKRVLTLES 359

DB 342 GYTREEIKKALTNQYNEVTATYLLGRKTEEGDRGAPGLALARVAPSDDTNGTSSSK 401

QY 360 GSGSRPAPPAGSV-----PLVPVPVPVPVPVPVPVPVPVPVPVPVPVPVPVP 402

DB 402 GSHNKGQRASSSTYHQRHSDFCGSPAPLH-PKESPTSTGDTLKEEMFPR 455

RESULT 12

Q8BWQ5 PRELIMINARY; PRT; 619 AA.

ID Q8BWQ5

AC Q8BWQ5

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Hypothetical eukaryotic protein kinase containing protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; Tissue=Liver;

RE MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK050312; BAC34182.1; -.

KW Hypothetical protein.

SQ SEQUENCE 619 AA; 69633 MW; FB5367F1DBA1E907 CRC64;

Query Match 15.4%; Score 344.5; DB 11; Length 619;

Best Local Similarity 30.7%; Pred. No. 1.8e-19;

Matches 103; Conservative 55; Mismatches 143; Indels 35; Gaps 10;

QY 6 PEPEPPSLTCCGP-----GTAPGPGAGVPLLTEDMQALTURTAAADSVTHKHYELRELK 61

DB 308 PEKESKRKLEEKPERPSPGRKPRP-----KGIISADVEKHDIYGVIGD 351

QY 62 GTYKGVLDLVYKGTGTAKLKFVNKSKTKLKNFL--REVSITNSLSSPFIIVFDVVE 119

DB 352 GNFAVTKECRHRETROAYAMKIIDKSLKGMVDSEILIIQSL-SHENIVKLHE-VYE 409

QY 120 TEDCYFAQYAPAGDLFDLIIPQVGLPDTVKRCVQQLGLADFMHGRVHRDIKPN 179

DB 410 TEAIVILMEYVOGGDLFDALIVENVKFEPERAAVMTDLCKALVHMDKNIVHRDKPEN 469

QY 180 VLFDFECRRV--KLADFGMTRRVGCRVKVSGTIPTTAPEVQOAGRADGLAVDTGVVW 237

DB 470 LLVQRNEDKSITLKLADFLGAKYVWRPIFTVCGTPTTYVAPEIL-SEKYGYLE----VDMW 524

QY 238 AFGVLFCVLTGNFWEAAS-GADAFEEFVRWQGRGLPGLPSQWRFFTEPALRMFORLL 296

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.M., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
CC Science 287:2185-2195(2000).
CL -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AE003798; AAF57652.1; -.
DR HSP; Q63450; 1A06.
DR FlyBase; FBgn0034376; CG15072.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 702 AA; 77356 MW; 441253E5A5484647D CRC64;

Query Match 16.1%; Score 361; DB 5; Length 702;
Best Local Similarity 29.1%; Pred. No. 9.7e-21;
Matches 109; Conservative 72; Mismatches 148; Indels 46; Gaps 15;

QY 18 GPGTAPGAGVPLLTEDMQALILRLAASDVTK-----HYELVRELKGTGYKVDLVVY 72
DB 8 GPAAP-PTSSTP-----QNYKVPSTSKISVDKLLRGVGYELEKTKGNFAVVKLATN 60
QY 73 KGTGTMALKFVNK---SKTKLKNFLREVSITNSLSPPFIKVDVVFVETDCYVFAQE 129
DB 61 IVYTKVAIKIDKTCNBEYVINKTFREIALKSL-RHPHILRYLYVM-ESQSYILVTE 118
QY 130 YAPAGDLFDIIPQVGLPDETVKRCVQVLGLALDFMGRQLVHRDIKPNVLLFDRECR 189
DB 119 YAPNGEITFDHLVANGRMKEPEARVFTQLVSAVHYCHRRGVVHRDLKAENVLL-DKD-MN 176
QY 190 VKLADFCWTRRV--GCRVKRVSGTIPVTAPEVCOAGRADGLAVDTGVDVWAFGLIFCVL 247
DB 177 IKLADFGSNHYEREGATIKTCWGSPPYAAPEVFGLEYDG-----PKSDIWSLGVVLYALV 232
QY 248 TGNFPPEAAAGADAFFEEFVRWQRGLPGLPSQWRRTTEPALRMFORLLALEPERRGPAK 307
DB 233 CGALPFDGKITLE--LKSRLVILGKFIPIFFNSQ-----ECEQLIRNMLVVEPDRRTIK 284
QY 308 EVFRLKHELTSELRPRSHARKPPGDRPPAPGLRLAPGLPKRTVLTSSGSGSPAP 367
DB 285 QI----IKHRLMSQSEMQEGERF--GDMSPGS-----GTVSKSASTSSLSGASDSP 331
QY 368 PAVGSVPLFVPPVPV 382
DB 332 FOLDSVNVVTHMLQLP 346

RESULT 9
Q960N7 PRELIMINARY; PRT; 512 AA.
AC Q960N7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE LD42896P.

GN GRP OR CG4711 OR CG17161.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RP Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CL -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AY051961; AAK93385.1; -.
DR FlyBase; FBgn0011598; grp_kinase.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 512 AA; 57833 MW; 504CAE3C0B642EAO CRC64;

Query Match 15.9%; Score 356.5; DB 5; Length 512;
Best Local Similarity 32.7%; Pred. No. 1.5e-20;
Matches 87; Conservative 49; Mismatches 111; Indels 19; Gaps 7;

QY 45 AASDPITKHYELVRELKGTGYKVDLVVYKGTGFMALKFVN-KSKTKLKNFLREVSITNS 103
DB 14 ATREFVEGVTLAQTLGEGAYGEVKLLINRQTGEAVAMKMWDLKKHPDAANSVRKEVCIOK 73
QY 104 LSSPFIKVDVVFETDCYVFAQYAPAGDLFDIIPQVGLPDETVKRCVQVLGLALD 163
DB 74 MLQDRHILFFGKRSQSVYEYIFL-EYAAAGGELFDRIEFDVGMFQHEAQRIFYTQLLSGLN 132
QY 164 FMHGRQLVRHDIKPNVLLFDRECRVVKLADFGWTRRVGCRVK-----RVSGTIPVTAPE 218
DB 133 YLHQGIARHDLKPNLLDDEHD--NVKISDFGMATMFRCKGKERLLDKCGTLPIVAPE 190
QY 219 VQO-AGRADGLAVDTGVDVWAFGLIFCVLTGTFNPWEAASGADAFEEFVRWQRGLPGL 277
DB 191 VLQKAYHAQ-----PADLWSCGVILVWMLAGELPMDQPS---TNCTEFTNWRDNDHWQL 241
QY 278 PSQWRRTTEPALRMFORLLALEPERR 303
DB 242 QTFWSKLDTLATSLRLKLLATSPGTR 267

RESULT 10
Q9C098 PRELIMINARY; PRT; 608 AA.
AC Q9C098
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA1765 (Fragment).
GN KIAA1765.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
DE The complete sequences of 100 new cDNA clones from brain which code

DR EMBL: AY071597; AAL49219.1; -
DR FlyBase; FBgn0031855; CGI1221.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 456 AA; 52418 MW; 272D73C6677C375B CRC64;

Query Match 19.6%; Score 438; DB 5; Length 456;
Best Local Similarity 34.1%; Pred. No. 3e-27;
Matches 113; Conservative 54; Mismatches 132; Indels 32; Gaps 8;

QY 5 CPPEPPRSRLTCCGPGTAPGGA-----GVPLLTEDMOALTLRTLAASD 48
DB 44 CSTPLTFTS-TSTGRLEVPGAASVRSSTYKKPKNDGGQIHLIPDVE-LPLMTFA--- 98
QY 49 VTKHYELVRELKGTGKGVLDVYKGTGKMAKLVFNKSKTKLKNFLREVSITNSLSSTP 108
DB 99 --DQYNIETKLAEGCFKILLCRHPTNTLVLKAVHAELTTIKFQKEFHYNYELSHH 156
QY 109 FLIKVDFVETECYVFAQYAPAGLFDLIPQVGLPDTVKRCVQOGLALDFMHGR 168
DB 157 HILSAYAVAFQTDYVYFAMHAPYGLASNGN-GLHNACKLISEQLSALLGFVHSK 215
QY 169 QLVHRDKPENLVLFDECRVRKLADFGMTRRVRCVRKVSQT---IPYAPVCCQAGR 224
DB 216 NLVHRDLKIENLVFTDFRVKLCDFGATTKGILLVHKVHTWTSVCPPEQLELIKNER 275
QY 225 ADGLAVDTGVDVMAFGLVIFCVLTGNFPWEAASGA-DAPFEFVRWQGRGLPGLPSQMR 283
DB 276 FQCLPVS---DSWQFGILLNLTGNFPWQSDWVDKQSYANFMKYEQRTTKVDPNFR 332
QY 284 FTPEALRMFORLLALEPERRGPAKEVFRFLK 314
DB 333 FSPRLMRCFRKLSYSHDPEDRCKITEVAKYMK 363

RESULT 7
Q90ZY5 ID Q90ZY5 PRELIMINARY; PRT; 503 AA.
AC Q90ZY5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein kinase Chk2.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RT "The zebrafish homolog of the human checkpoint kinase, Chk2."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF265346; AAK52419.1; -
DR InterPro; IPR000253; FHA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 503 AA; 56699 MW; EC01BF3D8E4CEC71 CRC64;

Query Match 16.5%; Score 370; DB 13; Length 503;
Best Local Similarity 30.6%; Pred. No. 1.2e-21;
Matches 107; Conservative 57; Mismatches 148; Indels 38; Gaps 13;

QY 32 LTEDMOALTLRTLAASDVTKHYELVRELKGTGKGVLDVYKGTGKMAKLVFNKSK--- 88
DB 178 MADDQANLPL-----EFSKKYHIAKIGTGVGCVKLAIEKETFKKVALTKINKHDFPS 231
QY 89 --TKLKNFLREVSIITNSLSSTPILIKVDFVVFTECYVFAQYAPAGLFDLIPQVGL 146
DB 232 IGATRNAERIEILKKI-DHPCLIKTED-FYTEDSYIVLYIEGGELEFGRIKAKKKL 289
QY 147 PETHVRCVQOGLALDFMHGROLVHRDIKPENVLL--FDRECRVRKLADFGMTR--RVG 202
DB 290 EEDIKLYFYQMLKAVEYLHNNGIIHRDLKPENVLLASHDDIC-LIKITDFNQSKILEES 348
QY 203 CRVRSVGTIPYAPVCCQAGRADGLAVDTGVVDVMAFGLVIFCVLTGNFPWEAASGADF 262
DB 349 SLMKTLTCLGTPYVAPEVFT--HASTVGYTKADVWSLGVLLFLICLGGYPFPFTECTTMSV 406
QY 263 FEFVVRWQGRGLPGLPSQMRRETEPALRMFORLLALEPERRGPAKEVFRFLKHELTSER 322
DB 407 REQI---NGEYFIISQMKVNSNEAKDLVKLLVDPQKRLSVEEA---LEHPWLKDDR 460
QY 323 RPSHRARKPPGDRPPAAGPLRLAEPGLKRTVLTESGSGS----RPAP 367
DB 461 MRQANQLNPG---AANQPMRPEATIRK---AQEGEGEPSSKRKPGP 503

RESULT 8
Q9V8L2 ID Q9V8L2 PRELIMINARY; PRT; 702 AA.
AC Q9V8L2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG15072 protein.
GN CG15072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKLEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers V.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwu C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

QY	289	LMFMORLLALPEPRGPAKEVFERFJ	313	
			:	
DB				

RESULT 6			
Q9VM90	PRELIMINARY;	PRT;	456 AA.
ID	Q9VM90		
CD	Q9VM90		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	CG11221 protein (RE51108p).		
GN	CG11221		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=BERKELEY;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pflogknoch C., Baldwin D.,		
RA	Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davis P.,		
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Hounck J.,		
RA	Hosini D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li C., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,		
RA	Falazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Swiskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Massarman D.A., Weinstein G.M., Weisenbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.-H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		
RT	"The genome sequence of Drosophila melanogaster.;"		
RL	Science 287:2185-2195 (2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RP	STRAIN=Berkeley;		
RA	Stapleton M., Brokstein P., Horg L., Agbayani A., Carlson J.,		
RA	Champe M., Chavez C., Dorsett V., Dresnek D., Fafan D., Frise E.,		
RA	George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,		
RA	Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,		
RA	Patel S., Phuanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,		
RA	Celniker S.;		
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.		
CC	-!- SIMILARITY: BELONGS TO THE SER/THER FAMILY OF PROTEIN KINASES.		
DR	EMBL; AF003614; AAF52433.1; -		

QY 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLRTLAASDVTKHYELVREL 60
DB 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLRTLAASDVTKHYELVREL 60
QY 61 KGYGKVDLVVYKGTGKTKALKFNKSKTKLKNFLREVSITNSLSPPFIKVFVDVVFET 120
DB 61 KGYGKVDLVVYKGTGKTKALKFNKSKTKLKNFLREVSITNSLSPPFIKVFVDVVFET 120
QY 121 EDCYVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
DB 121 EDCYVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
QY 181 LFDRECRVKLADFGMTTRVGRVCRVSGTIPYTAPEVQAGRADGLAVDTGVDVWAF 240
DB 181 LFDRECRVKLADFGMTTRVGRVCRVSGTIPYTAPEVQAGRADGLAVDTGVDVWAF 240
QY 241 VLIFCVLTGNFPWEAASGADAFEEFVVRWQGLPGLPSQWRFTPEPALRMFORLLALEP 300
DB 241 VLIFCVLTGNFPWEAASGADAFEEFVVRWQGLPGLPSQWRFTPEPALRMFORLLALEP 300
QY 301 ERGPAKEVFRFLKHELTSELRRPSSHRARKPGDPPAAAGPLRLLEAPGPKRTVLTESG 360
DB 301 ERGPAKEVFRFLKHELTSELRRPSSHRARKPGDPPAAAGPLRLLEAPGPKRTVLTESG 360
QY 361 SGSRPAPPVGSVPLPVVPVVPVVPVPEPGLAPGPGRTDGRADKSGQVVLATAI 420
DB 359 SGSRPSPSPVGPV-----VPVPVVPVVPVPEAGLAPPAPPGRTDGRADKSGQVVLATAI 413
QY 421 EICV 424
DB 414 EICV 417

RESULT 2

Q8QZX0 Q8QZX0 PRELIMINARY; PRT; 417 AA.
AC Q8QZX0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN SBK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; BC024114; AAH24114.1; -
DR EMBL; BC025837; AAH25837.1; -
DR EMBL; BC031759; AAH31759.1; -
DR MGD; MGI:2135937; SBK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN KINASE ST; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;

KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 417 AA; 45696 MW; C6A79A28EB6D2EC7 CRC64;
Query Match 92.6%; Score 2072.5; DB 11; Length 417;
Best Local Similarity 94.1%; Pred. No. 3.7e-160;
Matches 399; Conservative 3; Mismatches 15; Indels 7; Gaps 2;
QY 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLRTLAASDVTKHYELVREL 60
DB 1 MSVGCPEPPPSRLTCCGPGTAPGAGVPLLTEDMQALTTLRTLAASDVTKHYELVREL 60
QY 61 KGYGKVDLVVYKGTGKTKALKFNKSKTKLKNFLREVSITNSLSPPFIKVFVDVVFET 120
DB 61 KGYGKVDLVVYKGTGKTKALKFNKSKTKLKNFLREVSITNSLSPPFIKVFVDVVFET 120
QY 121 EDCYVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
DB 121 EDCYVFAQYAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALDFMHGRQLVHRDIKPEV 180
QY 181 LFDRECRVKLADFGMTTRVGRVCRVSGTIPYTAPEVQAGRADGLAVDTGVDVWAF 240
DB 181 LFDRECRVKLADFGMTTRVGRVCRVSGTIPYTAPEVQAGRADGLAVDTGVDVWAF 240
QY 241 VLIFCVLTGNFPWEAASGADAFEEFVVRWQGLPGLPSQWRFTPEPALRMFORLLALEP 300
DB 241 VLIFCVLTGNFPWEAASGADAFEEFVVRWQGLPGLPSQWRFTPEPALRMFORLLALEP 300
QY 301 ERGPAKEVFRFLKHELTSELRRPSSHRARKPGDPPAAAGPLRLLEAPGPKRTVLTESG 360
DB 301 ERGPAKEVFRFLKHELTSELRRPSSHRARKPGDPPAAAGPLRLLEAPGPKRTVLTESG 360
QY 361 SGSRPAPPVGSVPLPVVPVVPVVPVPEPGLAPGPGRTDGRADKSGQVVLATAI 420
DB 359 SGSRPSPSPVGPV-----VPVPVVPVVPVPEAGLAPPAPPGRTDGRADKSGQVVLATAI 413
QY 421 EICV 424
DB 414 EICV 417
RESULT 3
Q90ZY4 Q90ZY4 PRELIMINARY; PRT; 385 AA.
AC Q90ZY4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Protein kinase Bk146.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Chou C.-M., Leu J.-H., Huang C.-J.;
RT "Molecular cloning and characterization of a novel protein kinase expressed predominantly in the brain from zebrafish."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF265347; AAK52420.1; -
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN KINASE ST; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 385 AA; 43913 MW; 54BE4E6D3819EBD2 CRC64;
Query Match 45.1%; Score 1010; DB 13; Length 385;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:54:06 ; Search time 40 Seconds
(without alignments)
2735.358 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGPPEPRSLTCGPG.....GRADKSKGVLAIAICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2084.5	93.1	417	11 Q9Z335	Q9Z335 rattus norv
2	2072.5	92.6	417	11 Q8QZX0	Q8QZX0 mus musculu
3	1010.5	45.1	385	13 Q90ZV4	Q90ZV4 brachydanio
4	642.5	28.7	358	13 Q91758	Q91758 xenopus lae
5	452.5	20.2	551	5 Q8SZT8	Q8SZT8 drosophila
6	438	19.6	456	5 Q9VM90	Q9VM90 drosophila
7	370	16.5	503	13 Q90ZV5	Q90ZV5 brachydanio
8	361	16.1	702	5 Q9V8L2	Q9V8L2 drosophila
9	356.5	15.9	512	5 Q960N7	Q960N7 drosophila
10	346	15.5	608	4 Q9C098	Q9C098 homo sapien
11	345.5	15.4	752	11 Q8C1P4	Q8C1P4 mus musculu
12	344.5	15.4	619	11 Q8BQ05	Q8BQ05 mus musculu
13	344	15.4	512	10 P92958	P92958 arabisidopsis
14	343.5	15.3	508	10 Q9ZNT4	Q9ZNT4 oryza sativ
15	343.5	15.3	509	10 Q9ZRXJ	Q9ZRXJ oryza sativ
16	343	15.3	476	13 Q8AYC9	Q8AYC9 gallus gall

17	343	15.3	512	10 P92968	P92968 arabisidopsis
18	342.5	15.3	485	5 Q9N3L4	Q9N3L4 caenorhabdi
19	340.5	15.2	640	11 Q8COX8	Q8COX8 mus musculu
20	340.5	15.2	729	11 Q9JKE4	Q9JKE4 mus musculu
21	340.5	15.2	744	11 Q9JKE5	Q9JKE5 mus musculu
22	340	15.2	513	5 Q81661	Q81661 drosophila
23	339.5	15.2	716	4 Q8N399	Q8N399 homo sapien
24	336.5	15.0	535	10 Q8RWD2	Q8RWD2 arabisidopsis
25	336.5	15.0	752	4 Q8NG37	Q8NG37 homo sapien
26	336.5	15.0	797	11 Q8VHF0	Q8VHF0 rattus norv
27	336	15.0	504	10 P93113	P93113 cucumis sat
28	336	15.0	795	4 Q9P0L2	Q9P0L2 homo sapien
29	335	15.0	474	13 Q9Y1I8	Q9Y1I8 xenopus lae
30	334.5	14.9	517	13 Q98TW0	Q98TW0 xenopus lae
31	334.5	14.9	517	13 Q918V3	Q918V3 xenopus lae
32	334	14.9	778	4 Q8TDC2	Q8TDC2 homo sapien
33	334	14.9	793	11 Q08678	Q08678 rattus norv
34	334	14.9	795	11 Q8VHJ5	Q8VHJ5 mus musculu
35	333	14.9	476	11 Q9DON2	Q9DON2 mus musculu
36	332.5	14.8	504	10 Q41485	Q41485 solanum tub
37	332	14.8	510	10 Q40740	Q40740 oryza sativ
38	332	14.8	511	10 Q40544	Q40544 nicotiana t
39	332	14.8	513	10 Q40029	Q40029 hordeum vul
40	331.5	14.8	328	5 Q9NF25	Q9NF25 caenorhabdi
41	330.5	14.8	476	11 Q91ZN7	Q91ZN7 rattus norv
42	330.5	14.8	505	10 Q9ZRT1	Q9ZRT1 oryza sativ
43	330	14.7	514	10 Q9M726	Q9M726 lycopersico
44	329.5	14.7	291	10 Q9ZRA0	Q9ZRA0 arabisidopsis
45	329.5	14.7	494	10 Q9FLZ3	Q9FLZ3 arabisidopsis

ALIGNMENTS

RESULT 1

Q9Z335 ID Q9Z335 PRELIMINARY; PRT; 417 AA.
AC Q9Z335;
DT 01-MAY-1999 (TREMREL. 10, Created)
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Serine/threonine protein kinase (EC 2.7.1.-).
GN SBK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WistarST; TISSUE=Hippocampus;
RA Nara K.;
RT "Cloning and characterization of a novel serine/threonine protein
RT kinase gene expressed dominantly in developing brain.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB010154; BAB36362.1; --
DR HSBP; Q00534; 1B18.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR01245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SMC0220; S_TKc; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 417 AA; 45676 MW; C4D36608EE6EE208 CRC64;

Query Match 93.1%; Score 2084.5; DB 11; Length 417;
Best Local Similarity 94.6%; Pred. No. 3.9e-161;
Matches 401; Conservative 3; Mismatches 13; Indels 7; Gaps 2;

```
Db      11 SEVLKYNLGLTGLGTGFGVKAERNVTGQVAIKILNRRKXMTMEMEEKGNREIKIMR 70
QY      103 SLSS--SPFIKVFVETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGL 160
Db      71 LFIDFIHPHIIIVYEIETPKDIFV-VMEYCNGELLDYIENGRLQDEARRIFQOILA 129
QY      161 ALDFMHGROLVHRDIPENVLVDFDRECRVKVLDKADFGMTR--RVGCRVKVSGTIPYTAPE 218
Db      130 GVEYCHRMVVRDLDKPNELLDSR--YNVKADFGLSNVMDRGHFLKTKSCGLNAYAAPE 187
QY      219 VQAGRADGLAVDTGVDVWAFGLVFCVLITGNFPNFAASGADAFEEFVWRQGRPLPGLP 278
Db      188 II---SSKLYAGPEVDVWVSCGWLVALICGSVFED-----DDNIPSLFRKIKGGTYILP 238
QY      279 SOWRAFTEPALRMFORLLALEPERRGPAKE--VFRFLKHELTSELRRRPSHRARXPP 333
Db      239 SY---LSDSARDLIPKLLNIDPMKRITFHEIRVHPWFKNHL-----PCYLAVPFP 285

RESULT 15
G01025
serine/threonine protein kinase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Dec-1998
C:Accession: G01025
R:Navarro, E.
submitted to the EMBL Data Library, April 1996
A:Reference number: H00564
A:Accession: G01025
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-745 <NAV>
A:Cross-references: EMBL:X97630; NID:gl310674
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
F:18-271/Domain: protein kinase homology <KIN>

Query Match      14.4%; Score 322.5; DB 2; Length 745;
Best Local Similarity 30.2%; Pred. No. 1e-07;
Matches 80; Conservative 64; Mismatches 100; Indels 21; Gaps 9;

QY      52 HVELVRELKGYGKVDLVVYKGTGKVALKEVNSK---TKLKNFLREVSTNSLS SSP 108
Db      19 NVRLKLTIGKGNFAKVKLARHILTGKEVAVKIIDKTLNSSLQKLREVRIMKVL-NHP 77
QY      109 FLIKVFVVFETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHGR 168
Db      78 NIVKLFVI-EPEKTYLVMEYASGGEVFDYLVAHGRMKKEARAKFRQIVSAVCYCHQK 136
QY      169 QLVHRDIPENVLVDFDRECRVKVLDKADFGMTR--VGCVRKRVSGTIPYTAPEVCOAGRAD 226
Db      137 FIVHRDLKAENLL-LDAD-MNLIKADFGSNEFTFGNKLDTCGSPPYAAPLFOGKKYD 194
QY      227 GLAVDTGVDVWAFGLVFCVLITGNFPNFAASGADAFEEFVWRQGRPLPGLPSQWRFRTE 286
Db      195 G----PEVDVWVSLGVLITVSGSLPFDGQNLKE--LRERVLRGKYRIPFYS-----T 242
QY      287 PALRMFORLLALEPERRGPAKEVFR 311
Db      243 DCENLLKAKFLINPSKRGTEQIMK 267

Search completed: December 5, 2003, 09:58:17
Job time : 21 secs
```


RESULT 10

S60303
 serine/threonine-specific protein kinase (EC 2.7.1.1-- BKIN2 (version 1) - barley
 C:Species: Hordeum vulgare (barley)
 C:Date: 20-Jul-1996 #sequence_revision 27-Feb-1997 #text_change 11-Jun-1999
 C:Accession: S60303; S24578
 R:Halford, N.G.; Vicente-Carabajosa, J.; Sabelli, P.A.; Shewry, P.R.; Hannappel, U.; Krei
 Plant J. 2, 791-797, 1992
 A:Title: Molecular analyses of a barley multigene family homologous to the yeast protein
 A:Reference number: S60303; MUID:93258420; PMID:1302632
 A:Accession: S60303
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-513 <HAL>
 A:Cross-references: EMBL:X65606; NID:g18931; PIDN:CAA46556.1; PID:g18932
 C:Genetics: 62/1; 126/3; 187/3; 231/3; 293/3; 323/3; 351/3; 398/3; 476/3
 A:Introns: 62/1; 126/3; 187/3; 231/3; 293/3; 323/3; 351/3; 398/3; 476/3
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
 C:Superfamily: AMP-activated protein kinase; protein kinase homolog
 C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
 F:15-272/Domain: protein kinase homolog <KIN>
 F:23-31/Region: protein kinase homolog <KIN>
 F:46,65,143,145/Active site: Lys, Glu, Asp, Lys #status predicted
 F:148,152/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 14.8%; Score 332; DB 1; Length 513;
 Best Local Similarity 31.7%; Pred. No. 2.9e-08;
 Matches 96; Conservative 56; Mismatches 119; Indels 32; Gaps 10;
 C:47 SDVTKHYELRELKGTGKVDLVVYKGTGKTKALKFVNKSKTKL----KNFLREVSTIN 102
 Db 11 SEVLKNYNLGLTGLTGFDGVKVAHEKLTGQVAIKILRRKWEEMEEKANREIKMR 70

Qy 103 SLSS--SPFIKVFVETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGL 160
 Db 71 LFIDFHPHIIIRVYQVETPKDIFI-VMEYCNNGELLDVYIENGRLQDEARRIFQILA 129
 Qy 161 ALDFMHGRLVHRDIPKPNVLLFDRECRVKLADFGMTR--RVGCRVKVSGTIPYTAPE 218
 Db 130 GVEYCHRLMVHRDIPKPNVLLFDRECRVKLADFGMTR--RVGCRVKVSGTIPYTAPE 187
 Qy 219 VQAGRADGLAVDTGVDVYAFGLVFCVLTGNFPEAAAGADAFPEEFVWQGRGLPGLP 278
 Db 188 II-----SSKLAVAGREVWVSCGVILYALQGVDP-----DDNIPSLFRKIKGTYILP 238
 Qy 279 SQWRFTPEALRMFORLLALPERGPAKE--VRFKLHETLSELRRPSPHARKPPGDR 336
 Db 239 SY--LSDSARDLIPKLNIDPMKRTITIEIRVHPWFKNHL-----PCYLAVPPPYKA 288

Qy 337 PPA 339
 Db 289 PKA 291

RESULT 11

I48609
 probable serine/threonine-specific protein kinase (EC 2.7.1.1-- kem - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Mar-2000
 C:Accession: I48609; S31333
 R:Inglis, J.D.; Lee, M.; Hill, R.E.
 Mamm. Genome 4, 401-403, 1993
 A:Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.
 A:Reference number: I48609; MUID:93364122; PMID:8359177
 A:Accession: I48609
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-774 <INGL>
 A:Cross-references: EMBL:X70764; NID:g57919; PIDN:CAA50040.1; PID:g57920
 R:Inglis, J.D.; Lee, M.; Hill, R.E.
 submitted to the EMBL Data Library, January 1993
 A:Description: A novel protein kinase with homologues in yeast maps to mouse chromosome

A:Reference number: S31333

A:Accession: S31333

A:Molecule type: mRNA

A:Residues: 1-698, 'K', 700-702, 'GRGLPRAKASCTSGTCYAGMAHQATRT', 731-774 <ING2>
 A:Cross-references: EMBL:X70764

C:Genetics:

A:Gene: emk

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolc
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F:51-304/Domain: protein kinase homolog <KIN>
 F:59-67/Region: protein kinase ATP-binding motif

Query Match 14.8%; Score 330.5; DB 2; Length 774;
 Best Local Similarity 26.0%; Pred. No. 4.8e-08;
 Matches 99; Conservative 82; Mismatches 125; Indels 75; Gaps 15;

Qy 52 HVELVRELKGTGKVDLVVYKGTGKTKALKFVNKSK--TKLKNFLREVSTINSLSSP 108
 Db 52 NYRLTKTIGKGNFAKVKLARHILTGKEVAVKLIIDKTLQNLSSSLQKLFREVIMKVL-NHP 110

Qy 109 FIKVFVETEDCYVFAQYAPAGDLFDIIPQVGLPDTVKRCVQQLGLALDFMHGR 168
 Db 111 NIVKLFVI-EETKTLVMEYASGGEYFDLVVAHGRMKKEAKAKFQIVLHVQYCHQK 169

Qy 169 QLVHRDIPKPNVLLFDRECRVKLADFGMTR--VGCVRVKVSGTIPYTAPEVQAGRAD 226
 Db 170 FIVHRDLKAENLLD-DAD-MNIKIADFGFNEFTFGNKLDTFCGSPPYAAPLFGQKKID 227

Qy 227 GLAVDTGVDVYAFGLVFCVLTGNFPEAAAGADAFPEEFVWQGRGLPGLPSQWRRTTE 286
 Db 228 G----PEVDVWSGLVILYLVSGSLPFDGQNLKE--LRERVLRGKYRIPFVMS-----T 275

Qy 287 PALRMFORLLALPERGPAKEVFR-----FLKHELTSELRRRP----- 325
 Db 276 DCENLLKFLILNPSKGTLEQIMKDRWMNVGHDDDELKPVPEPLITTPGRDVRDVGVL 335

Qy 326 -----SHRAKPPGD-----RP-PAAGPLRLAEPG---LKR 353
 Db 336 HTEIQDSLVGQRVNEVWATYLLGYKSSSEPEGDITILKPRPSADLTNSSAPSPSHKVR 395

Qy 354 TVLTESGSGSPAPPAPVGSVP 374
 Db 396 SV-SANPKQRSSDOAVPAIP 415

RESULT 12

S5882
 protein kinase Cds1 (EC 2.7.1.1-- [validated] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 31-Dec-2000
 C:Accession: S5882; S71846; T41204; T52473
 R:Murakami, H.; Okayama, H.

A:Title: A kinase from fission yeast responsible for blocking mitosis in S phase.
 A:Reference number: S5882; MUID:95240713; PMID:7723827
 A:Accession: S5882

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-460 <MUR1>

A:Cross-references: EMBL:X85040; NID:g794146

R:Murakami, H.

submitted to the EMBL Data Library, March 1995

A:Reference number: S71846

A:Molecule type: mRNA

A:Residues: 1-60, 'G', 62-202, 'I', 204-460 <YUR2>

A:Cross-references: EMBL:X85040; NID:g794146; PIDN:CAA59410.1; PID:g794147

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.

submitted to the EMBL Data Library, July 1999

A:Reference number: 221978

A:Accession: T41204

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

```

Matches 95; Conservative 60; Mismatches 109; Indels 32; Gaps 12;

QY 47 SDVTKHYELVRLGKGTGYKVDLVVYKGTGKQALKFVNKSKTKLKNF-----LREVS I 100
Db 11 SEALKNYNLGRLTGIGISFGKVKIAEHKLTGHRVAIKILN--RRQMEMEMEKKAKREIKI 68
QY 101 TNSLSSSPFIKIVFDVVFETEDCYVPAQBYAPAGDLFDLIPQVGLPDTVK-RCVQQLG 159
Db 69 LR-LFTHPHIIRDYVIYTPDIYV-VMEYCKFGEFLDIYVEKGRLOEDRVAILRFSQII 126
QY 160 LALDFMHGSQLVHRDIKPNVLLFDRECRRVKLAIDFGMTTRY--GCRVVRVSGTTPYTPA 217
Db 127 SAVEYCHRNVAHRDLKPNLLDSK--YNVKLADFGLSNVMXHDGHFLKTSOGSPNYAAP 184
QY 218 ENVCQAGRADGLAVDGYDVAWGVLIFCVLTGNFPHEAASGADAFEEFVRVQGRLPGL 277
Db 185 EVI-SGK--LVAGRPVDVMSGCVILYALCGTLDPD-----DENIPNFKKIKGGIYTL 235
QY 278 PSOWRRFTPEALRMFORLALPEPRRGPAAKEVFRFLKHELTSLSRRRPSHRARKPP 333
Db 236 PS---HLSALARDLIPRLVVDPMKRIITREIRERQRFQI-----RLPYLAVPPP 283

RESULT 9*
A56009
C:serine/threonine-specific protein kinase (EC 2.7.1.-) NPK5 - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C:Accession: A56009
M:Murataka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994
A:Title: Characterization of tobacco protein kinase NPK5, a homolog of Saccharomyces cerevisiae.
A:Reference number: A56009; MUID:94217693; PMID:8164654
A:Accession: A56009
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-511 <MUR>
A:Cross-references: GB:D26602; NID:9496384; PIDN:BA05649.1; PID:9496385
C:Function:
C:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C:Superfamily: AMP-activated protein kinase; protein kinase homology
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F:17-271/domain: protein kinase homology <KIN>
F:25-33/Region: protein kinase ATP-binding motif
F:48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F:147,151/Binding site: magnesium (Asn, Asp) #status predicted

```

Query Match	14.8%;	Score 332;	DB 1;	Length 511;
Best Local Similarity	31.9%;	Pred. NO. 2.9e-08;		
Matches	92;	Conservative 56;	Mismatches 106;	Indels 34; Gaps 10;
Qy	51	KHYELVRELKGTGYKVDLWVYKGTGTMALKAFVNAKSTK-----LKNFIREVSIITNSLSS	106	
Db	17	RNYLKGKTIGISFGKVIKAETHLTGHKVAVKILNRRKTKNNMEMEEKVREIKILR-LFM	75	
Qy	107	SPRIIKVDFVWFETEDCYVFAEYAPAGBLFDIIPQVGLPEDTVKRCVQQQLGALDFMH	166	
Db	76	HPHIIILYEVETPSDIYV-VMEYKSGELFYIVKEKQLQDEDEARKFFQOIIISGVEYCH	134	
Qy	167	GROLVHRDIKPEENVLLFDECRVRVKLADFGMTR--RVGCRVRKVSUTIPTYTAPVCQAGR	224	
Db	135	RNMVVRDLKPENNLLDSK--WNVKIAFGLSNIMRDGHFLKTCGSPNYAAPEVI-SCK	191	
Qy	225	ADGLAVDTGDVNWAFGLIFCVLTGNFPWEASGADAFEEFVVRWQRGBLPGLPQSWRRF	284	
Db	192	--LYAGPEVDVMSCGVILYALLCGTLPEP-----DENIPNLFKKIKGGMISLPS---HL	240	
Qy	285	TEPALRMFORLLALEPERRGPAKEVERFLLKHELTSELRRRPSHRARP	332	
Db	241	SAGARDLIPMLITVDP-----MKRMTIPEIRHHPWFQALHP	276	

```

QY 258 GADAFEEFVRWQGBLGLPSQWRRTPEALRMFORLLALEPERRGPAKEVFR----- 311
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
258 LKE--LRERVLRGKYRIPFYS-----TDCENLLKRFVLNPIKRGITLEQIMKRWINA 309
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
312 -----FLKHEL-----TS 319
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
310 GHEEDELKPFVEPELDISDQKRIDIMVGMGYSOEIQESLSKMKYDEIATYLLLRKKS 369
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
320 EUR-----RRPSHRARP-----PDRPPAAGPLRLE---APGPL 351
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
370 EVRPSSDLNNSTQSPHHKVRQSVSSQKRRYSDHAGGIPSWAYPKRSQTSTADSDL 429
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 KRTVLTE-----SGSGSRPAPPAVGSVPLPVVPV 382
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
430 KEDGISRSKSTGSAVGKGIAPASPMIGNASNPNKADIP 468

RESULT 5
JC1446
serine/threonine-specific protein kinase (EC 2.7.1.1) AK21 - Arabidopsis thaliana
N:Alternate names: protein kinase SNF1 homolog
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 11-Jun-1999
C:Accession: JCI1446; S58266; S56334
R:LeGuen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.
Gene 120, 243-254, 1992
A:Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein
A:Reference number: JCI1446; MUID:93013041; PMID:1339373
A:Accession: JCI1446
A:Molecule type: DNA
A:Residues: 1-512 <LEG>
A:Cross-references: GB:M93023; NID:gl66599; PIDN:AAA32736.1; PID:gl66600
R:Thuemmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
submitted to the EMBL Data Library, May 1995
A:Description: Differential accumulation of the transcripts of 22 novel protein kinase
A:Reference number: S58256
A:Accession: S58266
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 144-198 <THU>
A:Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910
R:Thuemmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
Plant Mol Biol. 29, 551-565, 1995
A:Title: Differential accumulation of the transcripts of 22 novel protein kinase genes
A:Reference number: S66314; MUID:96123233; PMID:8534852
A:Accession: S66334
A:Molecule type: DNA
A:Residues: 144-198 <TH2>
A:Cross-references: EMBL:X86966; NID:g928909; PIDN:CAA60529.1; PID:g928910
C:Comment: This enzyme plays an important role in a signal transduction cascade regulati
C:Genetics:
A:Gene: Akin10; AK21
A:Introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C:Superfamily: AMP-activated protein kinase; protein kinase homolog
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F:17-271/Domain: protein kinase homolog <KIN>
F:25-33/Region: protein kinase ATP-binding motif
F:48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F:147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 15.0%; Score 336.5; DB 1; Length 512;
Best Local Similarity 30.3%; Pred. NO. 1.8e-08;
Matches 97; Conservative 59; Mismatches 111; Indels 53; Gaps 13;

QY 20 GPAGPGAGVPLLTEDMQALTTLTAASDVTKHYELVRLGKTYGKVDLVVYKGTGM 79
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 GSGTSGRSGV-----ESILPNYKLGRTLGIGSGFRVKIAEHALTGHKV 45

QY 80 ALKFVNKSKTK----LKNFLREVSTNLSGSPFIKVPDVFVETEDCVVFAQAPAGD 135
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 46 AIKILNRRKIKNMEEKVRREIKILR-LFMHPHIIRLYEVETPTDIYL-VMHEYNSGE 103
QY 136 LFDIIPQVGLPEDTVKRCVQQLGALDFMHCQQLVHRDIKPEENVLLFPRECRVRKLADF 195
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 104 LFDYIVKGRLODEARNFFQIISGVEYCHRMVVRDLKPENL--DSKC-NVKIADP 161
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 GMTR--RVCRKRVSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFGLVIFCVLTGNFWM 253
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 GLSNIMRDGHFLKTCGSPNYAAPEVI-SGK---LYAGEPVDVWSCGVLYALLCGTLDF 217
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 EAASGADAFEEFVRWQGBLGLPSQWRRTPEALRMFORLLALEPERRGPAKEVFR 312
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 D-----DENIPNLFKKIKGGIYTLPSH-----LSPGARDILPRMLVVDP----- 256
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 313 LXHELTSELRRRPSHRARP 332
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 MKRVTIPEIRQHPWFQAHLP 276

RESULT 6
TI0449
probable serine/threonine-specific protein kinase (EC 2.7.1.1) - cucumber
N:Alternate names: SNF1-related protein kinase
C:Species: Cucumis sativus (cucumber)
C>Date: 16-Jul-1999 #sequence revision 16-Jul-1999 #text_change 20-Jun-2000
C:Accession: TI0449
R:Gumpel, N.J.
submitted to the EMBL Data Library, December 1996
A:Reference number: Z17020
A:Accession: TI0449
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-504 <GUM>
A:Cross-references: EMBL:Y10036
A:Experimental source: cv. Masterpiece; cotyledon
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C:Superfamily: AMP-activated protein kinase; protein kinase homolog
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:6-260/Domain: protein kinase homolog <KIN>

Query Match 15.0%; Score 336; DB 2; Length 504;
Best Local Similarity 31.7%; Pred. No. 1.9e-08;
Matches 101; Conservative 57; Mismatches 107; Indels 54; Gaps 14;

QY 52 HYELVRELKGTGYKVDLVVYKGTGKALKFVNKSKTK---LKNFLREVSTNLSGSS 107
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 NYKLGKTLGIGSGFGVKIAEHALTGHKVAIKILNRKIKNLDMEKVRREIKILR-LPMH 65
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 108 PFIKVDVVFETEDCVVFAQAPAGDLFDIIPQVGLPEDTVKRCVQQLGALDFMHG 167
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 PHIRLYEVETPSDIYV-VMHEYKSGELFDYVEKGRLOQDEARNFFQIISGVEYCHR 124
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 168 ROLVHRDLKPEENVLLFPRECRVRKLADGMR--RVGCRVKVSGTIPYTAPEVCOAGRA 225
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 NKVVRDLKPENL--DSKC-NVKIADFGLSNIMRDGHFLKTCGSPNYAAPEVI-SGK- 180
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 DGLAVDTGVDVWAFGLVIFCVLTGNFFEEAASGADAFEEFVRWQGBLGLPSQ---- 280
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 --LYAGEPVDVWSCGVLYALLCGTLDFD-----DENTPNLFKKIKGGIYTLPSHLSGA 233
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 281 -----WRRFTPEALRM---FQ-----RLAL-EPERRGPAKEV-----F 310
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 234 RELIPSMVLVDPMKRITTEIRQHPWFQAHLPRLVLAAPPDPMQOAKKIDEDILQEVVKM 293
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 311 RFLKHELTSELRRRPSHRA 329
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 294 GDRNQLVSESRLRQNEA 312
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
T07415
probable serine/threonine-specific protein kinase (EC 2.7.1.1) PKIN1 - potato

```


C:\Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
 C:\Accession: T34074

R.Nelson, J.

submitted to the EMBL Data Library, November 1995

A:\Description: The sequence of *C. elegans* cosmid C01C4.

A:\Reference number: Z21472

A:\Accession: T34074

A:\Status: preliminary; translated from GB/EMBL/DBJ

A:\Molecule type: DNA

A:\Residues: 1-356 <N>

A:\Cross-references: EMBL:U41025; PIDN:AAA82347.1; CESP:C01C4.3

C:\Genetics:

A:\Gene: CESP:C01C4.3

A:\Introns: 18/3; 51/3; 109/1; 154/3; 205/1; 259/2; 288/3

C:\Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match 19.1%; Score 427.5; DB 2; Length 356;

Best Local Similarity 29.8%; Pred. No. 1.3e-12;

Matches 98; Conservative 75; Mismatches 136; Indels 19; Gaps 5;

48 DVTKHYELVRELKGTGYGVKVDLVVYKGTGTQKALKFVNKSKTKLNFLREVSTNLS 107

16 ILPNYKLGKTLGIGSGFKVIAEHVVTGHKVAIKLNRRKIKNMEMEKVREIKILR-L 74

8 DLEEVYTNKQLGTGRFGFKLAHKQSKQRIAKFFPRPQTKQADFVREYNYSFFLSH 67

108 PRLIKVDFVVFETED--CYVFAQEVAPAGDLFDIIP--QVGLPEDTVKRCVQQLGLALD 163

68 QNLIITYEGMFOSSDDTAYFFVQFCFRASLREAVEATNQAGIGRANTKKVFAAVLSAIE 127

164 FMHGRQLVHRDIKPNVLLFD--RECRVKLADFGMTTRVYGCVRKRVSGTIPYTAPEVCOA 222

128 FMEDENLVHNLKAENILFDANDISKVKTDFGLTRKVDTVTKYLEYNNYHAELCDT 187

223 GRADGLAVDTGVDVWAFGLVIFCVLTGTNFPNPPAAAGADAFEEFVRWQGRGLPGLSQWR 282

188 VVNEKLNVNKSTDIWALGIIFCYMKGKFPQKASIMCKPKYWEQWLKRNKPNALPKFN 247

283 RTEPALRMFORLLALEPERRGPAKEVREFL--KHELTSLRRRPSHRARKPGDPRPAA 341

248 PFEKALKLFXKSLTFRFDKRWTDKMRKCLAKELKLSVK-----RPEEDY 294

342 PLRLAEPGLKRTVLTSGSGSRPAPPA 369

295 YVMIDTASRQTATSSSGEPQDSAPA 322

RESULT 3

TS2633

serine/threonine-specific protein kinase (EC 2.7.1.1-) AKIN11 [validated] - Arabidopsis

N:\Alternate names: SNF1 protein kinase omolog AKIN11

C:\Species: Arabidopsis thaliana (mouse-ear cress)

C:\Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000

C:\Accession: T52633

R:\Bhalerao, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Machid

Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999

A:\Title: Regulatory interaction of PRL1 WD protein with Arabidopsis SNF1-like protein ki

A:\Reference number: Z25116; MUID:99238528; PMID:10220464

A:\Accession: T52633

A:\Status: preliminary; translated from GB/EMBL/DBJ

A:\Molecule type: mRNA

A:\Residues: 1-512 <BHA>

A:\Cross-references: EMBL:X99279; PIDN:CAA67671.1

A:\Experimental source: cultivar Columbia

C:\Genetics:

A:\Gene: AKIN11

C:\Function:

A:\Description: EC 2.7.1.1-; serine/threonine-specific protein kinase AKIN11 [validated, M

complements SNF1 mutations in yeast

C:\Superfamily: AMP-activated protein kinase; protein kinase homology

C:\Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

Query Match 15.4%; Score 344; DB 2; Length 512;

Best Local Similarity 28.5%; Pred. No. 8.5e-09;

Matches 120; Conservative 65; Mismatches 132; Indels 104; Gaps 20;

49 VTKHYELVRELKGTGYGVKVDLVVYKGTGTQKALKFVNKSKTK-----LNFLREVSTNLS 104

16 ILPNYKLGKTLGIGSGFKVIAEHVVTGHKVAIKLNRRKIKNMEMEKVREIKILR-L 74

105 SSSPFIKVFDFVVFETEDCYVFAQEVAPAGDLFDIIPQVGLPEDTVKRCVQQLGLALD 164

75 FMHGRQLVHRDIKPNVLLFD--RECRVKLADFGMTTR--RVGCRVXKVSQGTIPYTAPEVCOA 222

134 CHNNVVRHRLDKPENLL--DSRC--NIK--ADFGLSNWMMDGHFLTKTSCSPNYAAPEV-L 190

223 GRADGLAVDTGVDVWAFGLVIFCVLTGTNFPNPPAAAGADAFEEFVRWQGRGLPGLSQW- 281

191 GK--LYAGPEVDVWSCGVILYALLCG-LPPD-----DENIPNFKKIKGGIYTLPSHLS 242

282 -----RRTPEPALRM-----FO-----RLALE--PERRGPAKEV----- 309

243 SEARDLIPRLITVDPKRITITPEIRQHWFOTHLPRYLAVSPDPDTEQAKKINEIIVQEV 302

310 --FRELKHELTSELRRP-----SHRARKPGDPRPAPAGPFLRLAEPGLKKTIV 355

303 VMGEDRNQVLESNRNTQNDATVTVYLLDNRPVPSG-----YLS--EFQBT- 350

356 LTESGSGSRPAPPVAVGVSPLFPVVPVPPVPEPGLAQGP--PGRTDGRADSKGQ 413

351 -TDSSN-----PWRTEAGASPVGHWIPAHVDHYGLGARSQ 386

414 V 414

387 V 387

RESULT 4

S27966

probable serine/threonine-specific protein kinase (EC 2.7.1.1-) - human

N:\Alternate names: protein p78

C:\Species: Homo sapiens (man)

C:\Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 04-Mar-2000

R:\Maheshwari, K.K.; Som, S.; Parsa, I.

submitted to the EMBL Data Library, January 1992

A:\Description: Sequence of a cDNA encoding 78kD marker protein lost in chemically induce

A:\Reference number: S27966

A:\Accession: S27966

A:\Molecule type: mRNA

A:\Residues: 1-713 <MAH>

A:\Cross-references: EMBL:M80359; NID:gl89511; PIDN:AAA59991.1; PID:gl89512

C:\Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C:\Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:\54-307\Domain: protein kinase homology <KIN>

F:\62-70\Region: protein kinase ATP-binding motif

Query Match 15.2%; Score 340.5; DB 2; Length 713;

Best Local Similarity 24.4%; Pred. No. 1.6e-08;

Matches 112; Conservative 79; Mismatches 139; Indels 129; Gaps 17;

37 QALTILT-----LAASDVTKH-----YELVRELKGTGYGVKVDLVVYKGTGTQKALK 82

26 QEVTSRTSRSGARCNSTASCADQPHGTVRLTKTGKGNFAKVKARHLTLTGREVAIK 85

83 FVNKSK---TKLNFLREVSTNLSSSSPFIKVFVDFVVFETEDCYVFAQEVAPAGDLFDI 139

86 IIDKTQNLNPTSLQKLFREVIRMKIL--NEPNIVKLFVI--ETQKTLYLIMEVASGKGVFDY 143

140 IPPQVGLPEDTVKRCVQQLGLALDFMGRQVLRHDIKPNVLLFDRECRVRLADFGMTR 199

144 LVAGRMKEKEARSKFRQVSAVQYCHQKRIVHRDLKAENLL--DAD--WNIKIADFGFSN 201

200 R--VGCVRKRVSGTIPYTAPEVCOAGRADGLAVDTGVDVWAFGLVIFCVLTGNPWEAAS 257

202 EFTVGGKLDTCGSGPPYAAPELFOGKYDG-----PEVDVWVSLGVILYTLVSGSLPFDGQN 257

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2003, 09:54:36 ; Search time 20 Seconds

(without alignments)
2039.776 Million cell updates/sec

Title: US-10-016-985-2

Perfect score: 2239

Sequence: 1 MSVGPPEPPASLTCCGPG.....GRADSKGQVVLATAIEICV 424

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database : PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	642.5	28.7	358	1 S71887	serine/threonine-s
2	427.5	19.1	356	2 T34074	hypothetical prote
3	344	15.4	512	2 T52633	serine/threonine-s
4	340.5	15.2	713	2 S27966	probable serine/th
5	336.5	15.0	512	2 JCI446	serine/threonine-s
6	336	15.0	504	2 T10449	probable serine/th
7	332.5	14.9	504	2 T07415	probable serine/th
8	332	14.8	510	2 T04145	serine/threonine p
9	332	14.8	511	1 A56009	serine/threonine-s
10	332	14.8	513	1 S60303	serine/threonine-s
11	330.5	14.8	774	2 I48609	probable serine/th
12	328	14.6	460	2 S58882	protein kinase Cds
13	325	14.5	1032	2 D83637	serine/threonine p
14	323	14.4	513	1 S60304	serine/threonine-s
15	322.5	14.4	745	2 G01025	serine/threonine-s
16	318.5	14.2	473	1 S59941	serine/threonine p
17	318.5	14.2	1245	2 D86260	serine/threonine-s
18	317.5	14.2	576	2 T41587	probable carbon ca
19	317.5	14.2	602	2 S72513	FOG2 protein - yea
20	317.5	14.2	1398	2 T13741	hypothetical prote
21	316.5	14.1	1192	2 T18611	probable serine/th
22	316.5	14.1	1246	2 G89287	protein H39E23.1 l
23	315	14.1	520	2 G86414	probable protein k
24	315	14.1	552	1 A53621	hydroxymethylglut
25	315	14.1	560	2 A53600	phosphorylase kina
26	314.5	14.0	562	2 T29858	hypothetical prote
27	312	13.9	534	2 G88924	protein R02C2.1 li
28	311.5	13.9	1051	1 JW0051	serine/threonine-s
29	311	13.9	502	1 A41361	serine/threonine-s

30 309.5 13.8 735 2 A53300
31 309 13.8 552 1 S51025
32 309 13.8 633 1 A26030
33 309 13.8 798 2 JC7500
34 307 13.7 481 2 I49072
35 307 13.7 512 2 T07788
36 306 13.7 336 2 T21652
37 304 13.6 735 2 I51901
38 303 13.5 1081 2 S51899
39 302 13.5 1142 2 S59359
40 302 13.5 1558 2 T29253
41 301.5 13.5 440 2 T14736
42 300.5 13.4 740 2 I38556
43 299.5 13.4 580 2 T40939
44 298 13.3 856 2 T43631
45 295 13.2 435 2 E84707

ALIGNMENTS

RESULT 1

S71887

serine/threonine-specific kinase (EC 2.7.1.-), pk9.7 gastrula-specific - African clawed
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: S71887

R:Snake, A.M.; Smith, J.C.

EMBO J. 15, 4556-4565, 1996

A:Title: Regulation of embryonic cell division by a Xenopus gastrula-specific protein ki
A:Reference number: S71887; MUID:97042347; PMID:8887547

A:Accession: S71887

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-358 <SNA>

A:Cross-references: EMBL:X99406; NID:gl480369; PIDN:CAA67783.1; PID:gl480370

C:Genetics:

A:Gene: pk9.7

C:Function:

A:Description: involved in regulation of cell division

C:Superfamily: African clawed frog serine/threonine-specific kinase, gastrula-specific;

C:Keywords: phosphotransferase

F:9-274/Domain: protein kinase homology <KIN>

Query Match 28.7%; Score 642.5; DB 1; Length 358;
Best Local Similarity 46.5%; Pred. No. 3.9e-22;
Matches 125; Conservative 46; Mismatches 95; Indels 3; Gaps 2;
QY 47 SDVTKHVELVRELGLGTYGVKVDLVVYKGTCTGMALKFVNKSKTKLKNFLREVSTNLS 106
DB 5 SQIGDHVHKLGQGTFSHVLLARELISGHSVALKFVRKDRTRQAAFIHELNISIALSD 64

QY 107 SPFFIKVDVVFVEDCDYVFAQVAPADLDIIPPOVGLPEDTVKRCVQQLGLALDFMH 166

DB 65 YPGIITPTVETVDVFIQTQELAPATLSIITKEVIGIEEVVKRCVAVITALDFMH 124

QY 167 GRQLVHRDIKPNVLLDFRECRVKLADFGMTRRVGCRVKRVSGTIPYTAPEVCOAGRAD 226

DB 125 GRRLVHRDLKPDNLLMDKECVHKLCDFGTQTSVGLSPVSMSHIIPYMPFELCNLKNQ 184

QY 227 GLAVDTGVDVWAFGLIFCVLTGNFPWFAAGADAFEEFVRWQGR--LFGLSQWRPF 284

DB 185 LLVLDQSQMDIWSFGLLIFVALTYGFPWEEAVGHQNKQYQMFVHQNNRNVFA-PILNRF 243

QY 285 TEPALRMFORLLALEPERRRGAKEVERFL 313

DB 244 TQEWAMFFKLLSQKPSRSPDVTVLEFL 272

RESULT 2

T34074

hypothetical protein C01C4.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans